Linguistic Portfolios

Volume 9 Article 6

2020

Speech Style Variation of Vowels in Citation Form vs. Running Speech: Intelligibility Implications for AI-Enabled Devices

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

Jessica Krause
St. Cloud State University, jmkrause@stcloudstate.edu

Linguistic Portfolios–ISSN 2472-5102 –Volume 9, 2020 | 60

SPEECH STYLE VARIATION OF VOWELS IN CITATION FORM VS. RUNNING SPEECH: INTELLIGIBILITY IMPLICATIONS FOR AI-ENABLED DEVICES

ETTIEN KOFFI AND JESSICA KRAUSE[1]

ABSTRACT

Does the speech style used in producing vowels affect intelligibility? The question is worth asking because, in running speech, vowels have been found to play a greater role in intelligibility than consonants (Kewly-Port et al. 2007:2374 and Fogerty and Humes 2012:1500). Yaeger, a former member of Labov's research team (1975), claimed that the acoustic correlates of vowels vary greatly according to speech style. Ladefoged et al. (1976) disagreed and noted that the core acoustic correlates of vowels remain pretty much invariable, regardless of speech style. Evidence, either pro or con, has implications for how humans ought to interact with AI-enabled devices such as Siri, Alexa, Google Voice. We re-examine this unresolved issue on the basis of a comprehensive analysis of the 11 phonemic monophthong vowels of English produced by 22 speakers of American English (19 from Central Minnesota and 3 from Western Wisconsin). The acoustic correlates studied are F0, F1, F2, F3, F4, intensity, and duration. Just Noticeable Difference (JND) thresholds corresponding to each correlate are used to answer the research question. The findings discussed in this paper and the conclusions reached are based on 10,164 measured vowel tokens.

Keywords and phrases: Speech Style, Artificial Intelligence, Formant Analysis, Citation Form, Running Speech, Speech Intelligibility, AI-enabled devices, JND, Critical Band Theory.

1.0 Introduction

Lisker and Abramson (1964:407) remarked that the ultimate usefulness of measuring speech segments depends on how effectively it enables us to identify them in running speech. Furthermore, Kewly-Port et al. (2007:2374), Fogerty and Humes (2012:1500), and others have demonstrated that, as far as intelligibility is concerned, vowels are weightier than consonants when utterances occur in running speech. Therefore, there is a need to verify whether or not speech style affects intelligibility when people interact with their Artificial Intelligence (AI)-enabled devices such as Siri, Alexa, and Google Voice. This issue is more pressing now than ever before because, with the exponential growth of these "smart" devices, transactions that once required the physical presence of a person are now done remotely by speaking a sentence or two. Twenty-two participants who are native speakers of American English produced utterances containing the 11 monophthong phonemic vowels of English using two speech styles: Citation Form (CF) and Running Speech (RS). Assessing the intelligibility of their vowels in these two speech styles

[1]**Authorship responsibilities:** The first author assigned this topic to the second author who was enrolled in his acoustic phonetics course. Thereafter, he encouraged the second author and a classmate to present preliminary findings at the St. Cloud State University's Student Colloquium under the supervision of the first author. The second author collected data from 10 more participants. The first author also collected data from 12 more participants. The second author is recognized as such because of the data she collected and the measurements she did on 10 participants. The first author is solely responsible for writing this paper and for interpreting the results of the acoustic phonetic measurements. He bears full responsibility for any analytical or interpretive errors in this publication.

constitutes the main topic of this paper. Structurally, the paper is organized as follows. The first section describes the participants, data collection procedures, and the methodology. The second provides a very succinct review of speech styles. The third introduces the Critical Band Theory which provides us with the interpretive framework for the various acoustic correlates. Thereafter, each of the seven acoustic correlates (F0, F1, F2, F3, F4, intensity, duration) is discussed separately. The fourth installment highlights the implications of our findings for AI-enabled devices.

1.1 Experimental Designs

Talkers are unsure which speech style to adopt when interacting with their "smart" devices. Some over enunciate while others want to talk to these devices as naturally as they talk to human interlocutors. We investigate this issue by examining vowels that occur in two speech styles. In Experiment 1, 22 talkers (17 females and 5 males) were instructed to read each of the 11 monophthong phonemic vowels in Table 1 three times as naturally as possible. This reading corresponds to Citation Form (CF) speech style. These are the same words that have been used in various acoustic phonetic experiments since Peterson and Barney (1952).

Vowels	**heed**	**hid**	**hayed**	**head**	**had**	**hod**	**hawed**	**hoed**	**hood**	**who'd**	**hud**
Segments	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]

Table 1: Vowels in Citation Form

In Experiment 2, the same participants were also asked to read the following paragraph as naturally as possible. This corresponds to producing vowels in Running Speech (RS).

> Pl**ea**se call Stella. **A**sk her to bring these things w**i**th her from the store: Six g**oo**d spoons of fresh snow p**ea**s, five th**i**ck slabs of bl**ue** cheese, and m**ay**be a foot-long sandwich as a snack f**o**r her brother B**o**b. We als**o** need a small plastic snake, the little y**e**llow b**oo**k, a rubber d**u**ck, and a p**a**per I-p**a**d. She should not forget the d**o**g video game and the big toy fr**o**g for the k**i**ds. She must leave the faked gun at home but she may bring the ten sea turtles, the m**a**t that my mom b**ou**ght, and the black r**u**g. She can scoop these things into three r**e**d bags and two **o**ld backpacks. We will g**o** meet her, Sue, Jake, and Jenny Wednesday at the very l**a**st train station. The station is between the bus stop and the c**oo**kie store on Flag Str**ee**t. We m**u**st m**ee**t there 12 **o**'clock, for sure. The entrance **i**s at the **e**dge of the z**oo** in Zone 4 under the zebra sign. York's Treasure Bank is the tall building in the l**e**ft c**o**rner. She cannot miss it.

The vowels highlighted in red are those investigated in this paper. We note in passing that this text is a slightly longer version of the George Mason University's Speech Accent Archive text. As of November 15th 2019, 2883 participants have recorded themselves reading the original version.[2] The first author has felt the need to augment the text so as to account for the vowel /ʊ/ which was missing from the original and also to make sure that certain consonants appear in wider distributions.

[2] Information retrieved from http://accent.gmu.edu/ on November 21, 2019.

In each experiment, the vowels under consideration are isolated, annotated, and measured for F0, F1, F2, F3, F4, intensity, and duration. The procedures used to collect the measurements are illustrated by Figures 1 and 2:

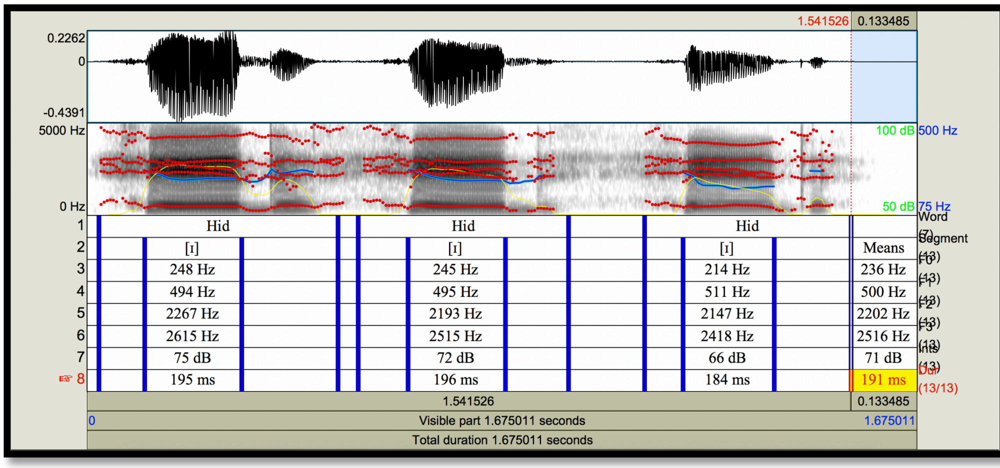

Figure 1: Words in Citation Form [Color Online][3]

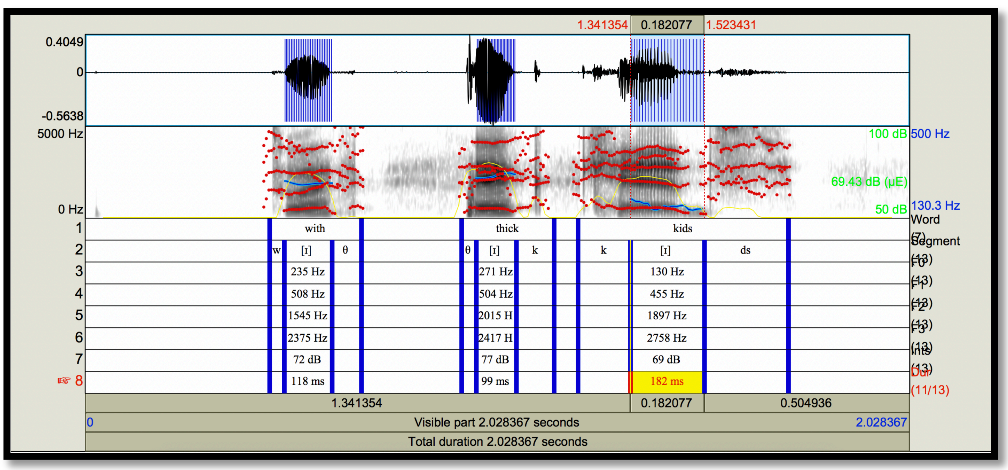

Figure 2: Words in Running Speech [Color Online]

All in all, the measurements yielded 10,164 tokens (11 vowels x 6 repetitions x 22 speakers x 7). The 17 female participants produced 7,854 tokes (3,927 vowel tokens in CF and 3,927 tokens in RS). The five male participants produced 2,310 (1,155 tokens CF and 1,155 tokens in RS). At the time of the study, 19 of the 22 participants were students at St. Cloud State University (SCSU) in St. Cloud Minnesota. They grew up in Central Minnesota. Three of the participants were from Western Wisconsin. They all signed a consent form approved by SCSU's the Institutional Review Board (IRB). Since the data produced by the female participants is far more substantive than those of their male counterparts, female data is displayed throughout the paper. Major insights are derived from female speech. Commentaries on male data are used as additional insights and/or confirmation of trends and patterns observed in female speech. The entirety of male data is provided in the appendices for ease of reference.

[3] All spectrographic and annotation analyses were carried out through Praat (Boersma and Weenink 2016).

1.2 Succinct Literature Review on Speech Style

Labov and his research team are among the first to have investigated the behavior of vowels in various speech styles by means of spectrographs. Labov (1972 :99) identified five such speech styles, as listed below:

1. Casual/vernacular speech style
2. Careful Speech style
3. Reading speech style
4. Word list speech style
5. Minimal Pair speech style

Labov (1972:76) displays vowels variants in different speech styles side by side in the same acoustic vowel space. Yaeger, a member of Labov's team (1975:3) used the same nomenclature and instrumental approach to investigate vowels in the same five speech styles. She measured the formant values of vowels and uncovered significant correlations between speech style and vowel formants, most notably F1 and F2. Ladefoged et al. (1976) were unconvinced by Yaeger's findings. So, they set out to investigate her claims by running their own experiments. They studied the vowels produced in the words <bee>, <bow>, <boy>, <bed>, <bad>, and <bud> in seven different speech styles, two more than in Yaeger's study. Nine male speakers of American English from California were recruited for their study. The researchers measured F1 and F2 frequencies of the vowels in these words. They also compared and contrasted the mean measurements of F1 and F2 against their standard deviations in all seven speech styles. The following are three of their most important findings:

1. "It may be seen that across all the styles of speech, we can estimate that two-thirds of all first formant frequencies are within 40 Hz of the mean, and two-thirds of all second formant frequencies are within about 90 Hz of the mean," (p. 230)
2. "In general, there is surprisingly little difference between styles," (p. 230)
3. "The more conversational styles of speech are usually (but not always) scattered around the points for the more formal styles. But there seems to be no systematic differences such that one can say that a given speaker in a certain style will have formant frequencies in a particular vowel systematically shifted in a certain direction away from the formant frequencies that occur when he is speaking in some other style," (p. 230)

Additional studies have been undertaken since. For example, Seung-Jae et al. (1994) investigated whether or not there was a difference between "Clear Speech" (CS) and CF. The former is defined as the speech style used in addressing somebody known to have a hearing impairment. Some additional findings, including the following, have come to light:

1. In CS, vowels are on average 40 to 60% longer than in CF (p.50)
2. Vowels are 3 to 5 dB louder in CS than in CF (p.43)
3. Vowels in CS tend to be more centralized (Gahl et al. 2012:3, 12-13).

It is important to note that these more recent studies have not concerned themselves with whether or not variations in speech style have any positive or negative impact on intelligibility. However, intelligibility is the main focus of this paper. The exponential growth in the use of AI-enabled automatic speech recognition devices has brought intelligibility to the forefront of speech style variation discussions. We will tackle this issue shortly, but first, we must introduce the theoretical framework under which acoustic measurements are interpreted in this study.

1.3 Theoretical Backbone

The previous studies employed different methodologies to arrive at their conclusions. Ladefoged et al. (1976) interpreted their data by considering arithmetic means and standard deviations. Seung-Jae et al. (1994) and Gahl et al. (2012:3) relied on various statistical instruments. In the current study, acoustic measurements are interpreted from the standpoint of the Critical Band Theory (CBT) and Just Noticeable Differences (JNDs). Since CBT is not well-known in mainstream linguistics, a quick overview is in order.

The theory emanated from Harvey Fletcher's groundbreaking psychoacoustic research at Bell Telephone Laboratory. Fletcher, a physicist, postulated on the basis of mathematical calculations that the basilar membrane is compartmentalized into frequency-sensitive areas called "critical bands". Von Bekesy, another physicist, proved clinically that Fletcher's theory was grounded in physiological reality. For his tireless effort and his genius in pioneering ways to probe into the inner ear, von Bekesy was awarded a Nobel Prize in Medicine in 1961. Figure 2 displays and summarizes pictorially the main tenets of CBT:

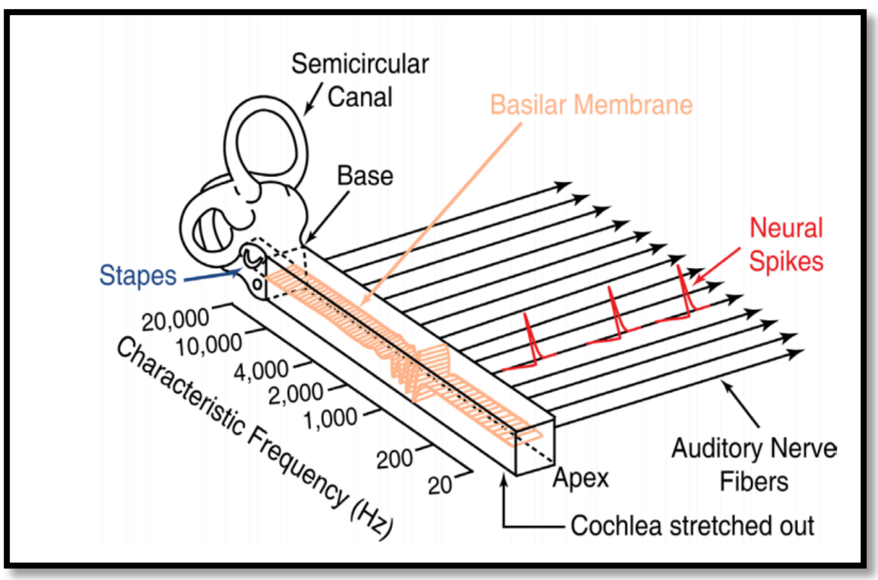

Figure 1: Audibility Range in the Frequency Domain [Color Online]
Sachs, M. B., Bruce, I. C., Miller, R. L., and Young, E. D. (2002). Biological basis of hearing aid design. *Annals of Biomedical Engineering* 30, 157-168. doi:10.1114/1.1458592. Reprinted by permission of © Biomedical Engineering Society.

A quick glance at the picture reveals the following. Humans perceive variations in frequency in designated areas of the basilar membrane. The 1/3 octave frequency bandwidth system is used in CBT because it is said to approximate as accurately as possible how humans process sounds in the frequency domain. The discovery of critical bands has led to other discoveries, including Just Noticeable Difference (JND) thresholds. Stevens (2000:225) explains that JNDs are correct responses elicited from experimental subjects. To qualify as a valid JND, at the very least 75% of the responses have to agree. Everest and Pohlmann (2015:23, 515) emphasize that JNDs are commonly used in physical and biomedical sciences to interpret data and establish degree of significance. In biomedicine, for example, JNDs are used to interpret myriads of measurements. Pre-established thresholds (i.e. JNDs) are used to gauge the significance of measurements. Experts also rely on JNDs to interpret the severity of physical phenomena such as hurricanes, tornadoes, flooding, and earthquakes. Universally accepted JNDs have been known and used in acoustics phonetics for nearly a century. They are, unfortunately, unknown to most linguists unfamiliar with acoustic phonetics. Koffi (2016) provides an introduction and a justification for using JND to interpret phonetic data. Suffice it to say that seven well-known JNDs are used in this paper to assess the significance of vowel variations in CF and RS speech styles.[4]

2.0 Vowel Duration in CF and RS

The JND for perceiving one segment as being longer than another is 10 ms. This JND is stated as follows:

JND in the Duration Domain

Vowels in CS are auditorily distinct from vowels in RS, and vice versa, if and only if the duration difference between them is ≥ 10 ms.

This JND was established by Fletcher as far back as 1929 (Fant 1960:233). Various other experiments have confirmed its reliability. For example, Crystal and House (1987:1555) studied 183,850 vowels in various speech styles. They found that vowels in RS with a slow tempo (108 ms) are 13 ms longer than vowels in RS with a fast tempo (95 ms). The durational difference between the two types of RS is therefore confirmed by the above-mentioned JND. If so, then it can also be used to determine whether or not there is a perceptual difference between words in CF and RS. Let's examine the data in Tables 2A through 2C to see if such is the case:

Duration	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	291	131	311	196	274	270	342	342	270	319	228
Speaker 2F	291	256	217	186	308	384	354	375	248	327	222
Speaker 3F	315	226	296	212	261	327	321	332	222	236	223
Speaker 4F	291	241	294	217	325	315	296	315	246	307	219
Speaker 5F	198	156	248	146	252	337	168	203	157	211	151
Speaker 6F	234	180	315	160	262	297	282	244	187	238	149
Speaker 7F	207	136	179	131	212	208	187	218	173	195	119
Speaker 8F	247	175	221	145	199	271	263	252	164	213	142
Speaker 9F	288	191	238	158	267	224	299	290	212	257	157
Speaker 10F	214	165	279	153	237	184	251	246	175	271	146
Speaker 15F	246	193	280	207	232	205	214	208	207	265	141
Speaker 16F	183	120	203	132	213	170	192	227	156	185	136
Speaker 17F	238	104	222	90	224	198	295	248	138	176	91
Mean	249	174	254	164	251	260	266	269	196	246	163
St. Deviation	42	46	44	37	37	67	60	55	41	50	44

Table 2A: Vowel Duration Vowels in Citation Form

[4] JNDs apply broadly. The formulations in this paper are tailored specifically for speech style analysis. However, these thresholds apply to any and all speech features of the same type regardless of speech style.

Duration	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	132	80	117	109	141	168	155	163	116	142	162
Speaker 2F	179	135	167	75	183	244	155	156	180	224	182
Speaker 3F	221	102	174	108	211	227	328	197	118	217	138
Speaker 4F	153	115	135	83	173	250	80	168	101	207	126
Speaker 5F	150	84	156	102	182	252	101	102	96	160	123
Speaker 6F	142	109	121	60	170	203	51	107	130	163	149
Speaker 7F	86	73	92	61	131	153	71	83	86	153	118
Speaker 8F	148	113	115	71	171	147	113	136	85	191	122
Speaker 9F	232	133	119	98	143	191	144	131	116	209	156
Speaker 10F	148	69	122	80	142	190	83	100	89	192	120
Speaker 15F	174	98	120	78	157	122	63	66	102	155	150
Speaker 16F	129	80	100	78	156	149	59	164	95	158	134
Speaker 17F	141	71	103	64	131	136	109	45	97	137	116
Mean	156	97	127	82	160	187	116	124	108	177	138
St. Deviation	38	22	25	17	23	45	72	44	25	30	20

Table 2B: Vowel Duration Vowels in Running Speech

Duration	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Vowels in CF	249	174	254	164	251	260	266	269	196	246	163
Vowels in RS	156	97	127	82	160	187	116	124	108	177	138
Difference	93	77	127	82	91	73	150	145	88	69	25

Table 2C: Summary of Duration Differences

Collectively, the vowels in CF lasted 226 ms, while those in RS lasted 133 ms. The duration difference between them is 93 ms. Consequently, we confirm that vowels in CF are 58.84% longer than those in RS. This finding is in agreement with the 40-60% range found in Seung-Jae et al. (1994:50). For male talkers (see Appendix 1), vowels in CF lasted 204 ms versus 125 ms in RS, that is, 61.27% longer. The difference between the two speech styles is 79 ms.

Clearly, a correlation exists between speech style and vowel duration. But does this difference matter for intelligibility? The answer is no because the human ear has an amazing capacity to integrate sounds (i.e., identify them accurately) in as little as 25 to 35 ms (Repp 1987:10, Everest and Pohlmann 2015:53). In other words, so long as a person is talking at a normal conversational speed, i.e., 120 to 200 words per minute (wpm), intelligibility is not compromised. AI-enabled devices can understand humans whose speech rate falls within this interval. Intelligibility is likely to be compromised only if the velocity of the speech is between 250 to 400 wpm or higher.[5] This speech style is equated with an auctioneer chant/speech. Since most people do not break into an auctioneer chant when talking to real human interlocutor, it is very unlikely that they would do so when talking to their AI-enabled devices. Consequently, if a person is talking normally, speech tempo would not prevent Siri, Alexa, or Google Voice from understanding their human interlocutors, assuming that instrumentation and software comply with

[5] Speech rate: https://www.write-out-loud.com/speech-rate.html. Retrieved on September 26, 2019.

ANSI (American National Standard Institute)[6] and ISO (International Standardization Organization)[7] specifications.

3.0 Speech Style Variation in the Frequency Domain

The human ear can perceive more frequencies than the mouth can produce. For example, a human with perfect hearing can perceive frequencies in the 20 to 20,000 Hz range (see Figure 1 above). However, for most ordinary speech styles, the frequencies that matter for intelligibility are situated on the 75 to 4,000 Hz frequency range. This bandwidth corresponds roughly to formants known as F0, F1, F2, F3, and F4. Everest and Pohlmann (2015:492) contend that F1, F2, and F3 are the most important formants, accounting for 75% of speech content. F0 plays a marginal role, if any, since English is not a tone language. F4 is hardly ever mentioned in intelligibility studies because it only relates to the size of the speaker's head (Ladefoged and Broadbent 1957:103). Whereas previous analyses of speech style variation have focused mainly on F1 and F2, in this paper we investigate all five formants because we do not want to leave any stone unturned.

3.1 Speech Style Variation and F0/Pitch

The JND for perceiving one segment as being higher in pitch than another is 1 Hz (Stevens (2000: 228, Lehiste 1970: 64, Gandour 1978: 57, Rabiner and Juang 1993:152, among others). When words are produced in isolation, vocal intensity is concentrated on said single lexical items whereas in running speech, vocal intensity is modulated in such a way that it is spread over the whole rhythmic group (Baken and Orlikoff 2000:110-1, 265). Consequently, everything being equal, one would expect the F0 of vowels in CF to be higher than their F0s in RS. Let's see if hypothesis is verified by the data:

JND in the F0 Domain

Vowels in CS are auditorily distinct from vowels in RS, and vice versa, if and only if the pitch difference between them is ≥ 1 Hz.

Vowels/F0	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	236	225	224	215	212	211	217	229	225	248	221
Speaker 9F	241	235	229	213	230	222	223	226	239	251	219
Speaker 11F	212	213	211	175	197	182	178	195	191	211	191
Speaker 12F	197	186	180	171	169	181	149	174	188	179	219
Speaker 13F	230	222	230	225	212	223	212	222	206	230	219
Speaker 14F	206	206	213	143	209	195	201	205	216	211	188
Mean	220	214	214	190	204	202	196	208	210	221	209
St. Deviation	17	17	18	32	20	19	28	21	19	27	15

Table 3A: F0 of Vowels in Citation Form

[6] Retrieved on November 23, 2019 https://www.ansi.org/.
[7] Retrieved on November 23, 2019 https://www.iso.org/home.html.

Vowels/F0	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	262	233	246	221	224	179	226	269	239	247	165
Speaker 9F	211	213	175	244	209	188	218	229	231	233	211
Speaker 11F	178	196	162	204	173	178	156	206	208	215	187
Speaker 12F	208	164	183	174	218	114	183	192	167	176	173
Speaker 13F	211	258	193	217	221	174	186	226	225	244	230
Speaker 14F	162	192	222	196	153	176	176	212	219	214	210
Mean	205	209	196	209	199	168	190	222	214	221	196
St. Deviation	34	33	31	23	29	26	26	26	25	26	25

Table 3B: F0 of Vowels in Running Speech

F0	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	220	214	214	190	204	202	196	208	210	221	209
Running Speech	205	209	196	209	199	168	190	222	214	221	196
Difference	15	6	24	19	5	34	6	14	4	0	13

Table 3C: Summary of F0 Differences

The arithmetic means in Tables 3A through 3C confirm the hypothesis. In female speech, the vocal folds vibrates higher in CF (208 Hz) than in RS (202 Hz). The difference of 8 Hz is perceptually salient. The same is true for male speech where F0 in CF is 109 Hz compared with 107 Hz in RS (see Appendix 2). The difference of 2 Hz is also perceptually salient. However, since English is not a tone language, differences in pitch between CF and RS have no impact whatsoever on intelligibility. This conclusion is further corroborated by Lieberman and Ryant (2016:77). They extracted F0 correlates from a wide variety of speech styles and arrived at the following conclusion:

> Although some studies have found pitch-range differences between read and spontaneous speech, it's intuitively clear that a speaker can use a wider or narrower pitch range in either situation. And this is what we found …. This doesn't mean that the pitch range dimension is not relevant or useful in general – it's clearly a significant aspect of prosodic style – but it's not relevant or useful to the specific task of distinguishing spontaneous speech from reading.

3.2 Speech Style Variation and F1

F1 correlates with tongue height in the production of vowels. Higher vowels have numerically lower measurements than lower vowels. This is known as inverse proportionality. Let's take the mean measurements of [i] and [æ] as examples. Because [i] is a high vowel, its F1 measurement (370 Hz) is smaller than [æ] (837 Hz), which is a low vowel. F1 is universally recognized as having the lion's share of vowel intelligibility because, as noted by Ladefoged and Johnson (2015:202), it alone contains 80% of the acoustic energy found in vowels. A similar statement and explanations are found in Kent and Read (2002:33,132-4).

The JND for the optimal perception of F1 has been studied by many researchers and used in many acoustic phonetic analyses, including Mermelstein (1978:578), Hawks (1994:1079), Rabiner and Juang (1993:152), and Labov et al. 2013:43), among others. The consensus is that for an optimal perception, the JND should be ≥ 60, as stated below:

JND in the F1 Domain

> Vowels in CS are auditorily distinct from vowels in RS, and vice versa, if and only if the F1 difference between them is ≥ 60 Hz.

If we find that the F1 of vowels in CF and RS differ by this magnitude, then speech style makes a difference in intelligibility. Otherwise, it does not. Let's query the data to see what we get.

Vowels/F1	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	359	500	452	739	896	896	839	550	610	383	696
Speaker 2F	368	506	503	851	945	858	825	516	553	429	773
Speaker 3F	293	437	450	723	771	811	667	490	470	373	670
Speaker 4F	378	524	464	731	853	798	813	516	574	407	746
Speaker 5F	393	325	399	466	845	814	807	499	535	400	687
Speaker 6F	362	476	476	738	927	920	852	575	506	441	745
Speaker 7F	361	550	454	759	916	905	859	619	577	455	730
Speaker 8F	450	491	471	622	800	836	739	543	533	426	585
Speaker 9F	361	500	481	702	922	843	820	598	543	440	703
Speaker 10F	337	427	435	628	807	752	691	531	483	379	603
Speaker 11F	354	473	426	703	781	786	774	429	445	415	706
Speaker 12F	464	472	469	627	745	718	703	530	492	406	556
Speaker 13F	344	439	502	596	730	783	799	490	432	428	550
Speaker 14F	338	430	474	704	847	899	860	536	487	400	700
Speaker 15F	363	475	410	611	751	824	769	535	549	402	666
Speaker 16F	396	535	495	643	871	702	740	487	491	403	607
Speaker 17F	384	446	413	578	830	830	856	506	487	417	686
Mean	370	470	457	671	837	822	789	526	515	412	671
St. Deviation	40	52	31	88	68	63	62	44	48	22	67

Table 4A: F1 of Vowels in Citation Form

Vowels/F1	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	405	544	489	734	924	845	692	614	648	417	790
Speaker 2F	406	529	487	722	837	797	557	551	573	421	721
Speaker 3F	361	376	421	604	892	878	505	525	603	365	657
Speaker 4F	412	547	464	689	937	843	547	526	571	468	783
Speaker 5F	415	493	436	673	827	742	526	473	525	487	657
Speaker 6F	379	503	465	732	885	738	637	414	551	508	759
Speaker 7F	393	575	570	760	952	897	640	637	632	476	804
Speaker 8F	398	351	443	523	610	768	506	457	572	398	691
Speaker 9F	417	489	424	394	846	776	614	562	527	421	604
Speaker 10F	396	467	447	643	791	669	625	599	512	417	622
Speaker 11F	369	450	427	695	771	788	631	596	489	401	640
Speaker 12F	465	496	484	583	767	741	594	510	494	391	575
Speaker 13F	420	425	415	600	753	743	507	508	495	411	562
Speaker 14F	410	495	482	676	824	779	618	531	550	461	747
Speaker 15F	396	489	467	597	762	758	495	456	512	372	588
Speaker 16F	426	535	501	606	856	769	587	532	531	430	637
Speaker 17F	411	493	435	587	732	733	539	467	527	430	630
Mean	404	485	462	636	821	780	577	526	547	427	674
St. Deviation	23	58	38	90	86	57	59	61	46	40	78

Table 4B: F1of Vowels in Running Speech

F1	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	370	470	457	671	837	822	789	526	515	412	671
Running Speech	404	485	462	636	821	780	577	526	547	427	674
Difference	34	15	5	35	16	42	245[8]	0	32	15	3

Table 4C: Summary of F1 Differences

The overall F1 mean of vowels in CF is 594 Hz compared with 576 Hz in RS. The acoustic difference between them is 18 Hz. This is below the threshold of 60 Hz in the JND. The fact that the acoustic difference is below 20 Hz is also very important. Numerous studies, including Kent and Read (2002:110) and Thomas (2011:56) report that in the F1 frequency band, humans cannot detect frequencies that are ≤ 20 Hz. In fact, in male speech, the F1s of vowels in CF and RS are identical, that is, 479 Hz each (see Appendix 3). This shows that the F1 of vowels does not vary even if the speech style varies. Our findings are identical with Ladefoged et al. (1976:230). The following statement from their study sums up the lack of variability beautifully, "In general, there is surprisingly little difference between styles."

3.3 Speech Style Variation and F2

In vowel articulation, F2 correlates with the horizontal movement of the tongue. There are three articulation targets along the F2 continuum: front, central, and back. Front vowels have higher F2 measurements than central vowels, which in turn have higher F2 values than back vowels. Mermelstein (1978:578) found that the JND for F2 was 171 Hz. Rabiner and Juang (1993:152) report a JND of 158 Hz, and Scharf (1961:215) has 200 Hz. The latter is in keeping with center of frequency estimations based on the 1/3 octave system. Since this estimation is universally accepted for the manufacturing and engineering of audio products, we use it in all intelligibility analyses (Pope 1998:1346). This leads to the following threshold formulation:

JND in the F2 Domain

Vowels in CF are auditorily distinct from vowels RS, and vice versa, if and only if the F2 difference between them is ≥ 200 Hz.

Kent and Read (2002:111) indicate that F2 is sensitive to dialectal variations. For this reason, we restate again that 2 of the 22 participants are from Wisconsin. The rest are from Minnesota, specifically from Central Minnesota. Tables 5A through 5C display their data.

Vowels/F2	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	2561	1936	2223	1829	1591	1415	1397	1218	1563	1145	1513
Speaker 2F	2536	1964	2321	1698	1601	1378	1118	1008	1517	1075	1510
Speaker 3F	2688	2413	2552	2255	2118	1531	1692	1389	1834	1304	1854
Speaker 4F	2845	2099	2645	2096	1706	1289	1308	947	1384	958	1339
Speaker 5F	2658	2246	2446	1935	1898	1568	1508	1012	1746	991	1741
Speaker 6F	2847	2428	2593	2079	1839	1361	1314	1109	1706	1248	1352

[8] The huge difference between CF and RS is a further indication that this vowel is in a great state of flux in Central Minnesota English. It has merged with [ɑ] before stop consonants but it still retains its phonemic identity before liquids (Koffi 2013:12).

Speaker 7F	2880	2455	2713	1992	1922	1514	1473	1207	1621	1558	1811
Speaker 8F	2669	1663	2348	1863	1868	1424	1382	1063	1461	1306	1543
Speaker 9F	2587	2202	2359	1927	1740	1329	1294	1146	1562	1299	1516
Speaker 10F	2842	2335	2672	2023	1840	1295	1284	1390	1584	1583	1602
Speaker 11F	2833	1817	2043	1565	1758	1423	1304	957	1578	1136	1496
Speaker 12F	2046	2012	2034	1542	1515	1166	1034	1104	1249	1338	1394
Speaker 13F	2684	2278	2518	2158	2013	1532	1442	1102	1618	1234	1646
Speaker 14F	2848	2158	2257	1988	1851	1531	1438	1207	1526	1229	1670
Speaker 15F	2539	2199	2472	1922	1817	1537	1516	1101	1627	974	1581
Speaker 16F	2703	2164	2468	2093	1942	1162	1301	1054	1309	1030	1606
Speaker 17F	2657	2216	2490	2178	1856	1328	1290	1030	1333	1207	1730
Mean	2671	2152	2420	1949	1816	1399	1358	1120	1542	1212	1582
St. Deviation	253	93	89	69	129	92	63	79	193	79	141

Table 5A: F2 of Vowels in Citation Form

Vowels/F2	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Spch	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	2333	1886	2181	1590	1573	1406	1426	1388	1554	1718	1579
Speaker 2F	2473	1423	2391	1524	1397	1237	1013	1275	1392	1353	1461
Speaker 3F	2679	2008	2577	1883	1944	1575	1393	1295	1650	1698	1749
Speaker 4F	2501	2125	2612	1597	1776	1302	1332	1229	1362	1416	1606
Speaker 5F	2548	2025	2408	1369	1816	1480	1314	1168	1545	1042	1710
Speaker 6F	2566	1930	2552	1824	1721	1337	1531	1063	1478	1788	1575
Speaker 7F	2574	1996	2489	1734	1820	1511	1675	1617	1455	1856	1730
Speaker 8F	2260	1973	2386	1748	1726	1398	1294	1369	1571	1628	1612
Speaker 9F	2359	1819	2111	1866	1645	1340	1367	1358	1679	1620	1585
Speaker 10F	2322	2061	2309	1662	1626	1251	1340	1172	1562	1671	1491
Speaker 11F	2487	1990	2453	1733	1899	1459	1365	1235	1303	1699	1594
Speaker 12F	2160	1897	2074	1416	1757	1197	1102	1209	1358	1475	1473
Speaker 13F	2469	2134	2476	1688	1954	1435	1320	1310	1505	1633	1526
Speaker 14F	2451	1995	2333	1618	1684	1352	1265	1253	1319	1586	1487
Speaker 15F	2453	2045	2399	1625	1740	1492	1350	1195	1478	1315	1625
Speaker 16F	2324	1777	2290	1646	1665	1381	1414	1086	1329	1377	1473
Speaker 17F	2433	2047	2326	1715	1693	1355	1038	1366	1266	1598	1534
Mean	2434	1948	2374	1661	1731	1382	1325	1269	1459	1557	1577
St. Deviation	295	114	65	130	81	147	335	268	120	303	101

Table 5B: F2 of Vowels in Running Speech

Vowels/F2	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	2671	2152	2420	1949	1816	1399	1358	1120	1542	1212	1582
Running Speech	2434	1948	2374	1661	1731	1382	1325	1269	1459	1557	1577
Difference	237	204	46	288	85	17	33	149	83	345	5

Table 5C: Summary of F2 Differences

The overall F2 mean for female speakers in CF and RS are respectively 1747 Hz and 1701 Hz. Since the F2 difference between them is only 46 Hz, this is further proof that the female participants in our study did not change the horizontal movement of their tongue in any appreciable way. Male participants also did not change their tongue movement much because their F2 values in CF and RS are respectively 1536 Hz and 1497 (see Appendix 4). The acoustic difference of 39 Hz is well below the JND threshold. Again, here as in the previous section, our findings are in perfect agreement with Ladefoged et al. (1976:230). In other words, speech style does not have any appreciable impact on F2.

3.4 Speech Style Variation and F3

F3 is commonly correlated with lip movement. Delattre (1951:873) also correlates it more with the raising or lowering of the velum. Bradley (2018:382-3), on the other hand, correlates it with the raising and the lowering of the larynx. All three correlations yield the same acoustic results. When the velum or larynx are raised, the lips tend to be spread. This leads to higher F3 measurements. When the velum or the larynx are lowered, the lips tend to be rounded. This articulatory gesture causes F3 to be lower. The JND of F3 on the Critical Bands system in Scharf (1961:215), Fastl and Zwicker (2007:235-5, and Everest and Pohlmann (2015:13) is taken to be at 400 Hz. Rabiner and Juang (1993:152) report a slightly smaller JND of 355 Hz. Everything being equal, the difference of 45 Hz does not amount to much. Therefore, we go with the JND threshold that is based on critical band estimations given the aforementioned reasons:

JND in the F3 Domain

Vowels in CF are auditorily distinct from vowels in RS, and vice versa, if and only if the F3 difference between them is ≥ 400 Hz.

Let's examine the data in Tables 6A through 6C to see if speech style has any impact on F3.

Vowels/F3	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	2961	2852	2714	2773	2046	2418	2633	2833	2787	2933	2700
Speaker 2F	2968	2170	2654	2273	2264	2619	2771	2728	2599	2690	2635
Speaker 3F	3396	3017	3159	2996	2856	2679	2799	3329	3136	3023	2978
Speaker 4F	3585	2088	3052	2832	2435	3109	3036	3293	3096	2858	2673
Speaker 5F	2986	2509	2862	2453	2359	2577	2552	2690	2683	2787	2691
Speaker 6F	3534	3126	3069	2876	2548	2533	2540	2736	2731	3061	2826
Speaker 7F	3425	3069	3154	2754	2350	2798	2830	3066	2985	3012	2795
Speaker 8F	2998	2859	2583	2885	2668	2960	2889	2798	2737	2697	2835
Speaker 9F	2831	2516	2679	2522	2590	2775	2781	2728	2667	2630	2708
Speaker 10F	3373	2894	3143	2791	2746	2485	2495	2477	2593	2616	2592
Speaker 11F	3372	2644	2804	2340	2410	2409	2278	2342	2386	2845	2518
Speaker 12F	3094	2979	2927	2913	2742	2773	2643	2848	2731	2600	2771
Speaker 13F	3251	3061	3099	2953	2793	2928	2887	2847	2845	2825	2858
Speaker 14F	3402	2586	2682	1614	2383	2675	2679	2739	2852	2834	2891
Speaker 15F	2926	2871	2864	2680	2671	2706	2631	2988	2590	2948	2715
Speaker 16F	3064	2963	2953	2934	2788	2528	2836	2832	2761	2666	2793
Speaker 17F	2974	2496	2678	2575	2348	2781	2501	2772	2823	2971	2689
Mean	3184	2747	2886	2656	2529	2691	2693	2826	2764	2823	2745
St. Deviation	242	311	199	346	226	194	187	246	187	152	114

Table 6A: F3 of Vowels in Citation Form

Vowels/F3	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Spch	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	2926	2895	2514	2532	2444	2504	2046	2886	3017	2734	2579
Speaker 2F	2864	2650	2845	2272	2371	2634	2353	2827	2766	2679	2486
Speaker 3F	3357	3103	3202	2817	2968	2818	2511	3259	3335	3200	2970
Speaker 4F	3148	2081	3024	2332	2386	2647	2536	2953	3098	2954	2333
Speaker 5F	3039	2544	2936	2452	2572	2538	2323	2710	2738	2893	2415
Speaker 6F	3172	2780	2999	2750	2713	2507	2552	3190	2876	3152	2373
Speaker 7F	3061	2971	3037	2603	2551	2818	2545	3009	2716	2986	2665
Speaker 8F	2801	2797	2804	2518	2416	2816	2535	3030	2950	1757	2749
Speaker 9F	2599	2516	2560	2627	2493	2382	1952	2696	2594	2678	2468
Speaker 10F	2824	2792	2746	2735	2569	2238	2398	2717	2566	2682	2364
Speaker 11F	3034	2781	2862	2423	2697	2240	2442	2848	2685	3004	2334
Speaker 12F	2815	2969	2653	2649	2809	2538	2553	2747	2718	2979	2657
Speaker 13F	2943	2958	3037	2931	2942	2718	2692	2866	2928	2855	2829
Speaker 14F	2869	2709	2666	2606	2719	2535	2198	2889	2842	2715	2555
Speaker 15F	2882	2777	2799	1914	2775	2582	2462	2976	2854	2676	2454
Speaker 16F	2907	2876	2786	2508	2597	2560	2307	2980	2838	2339	2566
Speaker 17F	3324	2648	2782	2462	2321	2705	2359	3042	2890	3230	2720
Mean	2974	2755	2838	2537	2608	2575	2397	2919	2847	2794	2559
St. Deviation	195	233	184	233	195	176	192	161	188	352	184

Table 6B: F3 of Vowels in Running Speech

Vowels/F3	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	3184	2747	2886	2656	2529	2691	2693	2826	2764	2823	2745
Running Speech	2974	2755	2838	2537	2608	2575	2397	2919	2847	2794	2559
Difference	210	8	48	119	79	116	296	93	83	29	186

Table 6C: Summary of F3 Differences

The overall arithmetic mean of F3 in CF is 2776 Hz. In RS, it is 2709 Hz. The acoustic difference between the two styles of speech is only 67 Hz. In male speech, the means are respectively 2603 Hz in CF, and 2589 in RS (see Appendix 5). The acoustic difference between the two styles of speech is only 14 Hz. These measurements align with what we have found so far, namely that variations in speech style have no corresponding effects on the intrinsic acoustic characteristics of vowels.

3.5 Speech Style Variation and F4

It has been noted that F4 is a formant that carries practically zero linguistic information. The prevailing consensus is that it correlates with the size of the head of the speaker:

> No simple technique will enable one to average out the individual characteristics so that a formant plot will show only the phonetic qualities of the vowels. One way to deal with this problem is probably to regard the average frequency of the fourth formant as an indicator of the individual's head size and then express the values of the other formants as percentages of the mean fourth formant frequency. But this possibility is not open when the fourth formant frequencies have not been reported for the sets of the vowels being compared (Ladefoged 2006:205-6).

It is, therefore, very unlikely that F4 will vary in response to speech style. However, let's subject this formant to the same JND test as done previously. According to critical band estimation, the JND of F4 is set at 600 Hz (Stevens 2000: 154, 300, Fastl and Zwicker 2007:235-5, Everest and Pohlmann 2015:13), and Scharf (1961:215). Not surprisingly, the JND of F4 is at 480 Hz in Rabiner and Juang (1993:186). As we have seen in previous cases, the JNDs proposed by Rabiner and Juang are slightly less than those reported by other scholars. Even so, the differences are not enough to affect the results of any intelligibility analysis. Therefore, we go with the majority view.

JND in the F4 Domain

> Vowels in CF are auditorily distinct from vowels in RS, and vice versa, if and only if the F4 difference between them is ≥ 600 Hz.

Let's examine the data in Tables 7A through 7C to determine the potential impact of F4 on speech style.

Vowels/F4	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 9F	3209	2939	2913	2967	3199	3497	3434	3485	3535	3663	2973
Speaker 11F	4271	3718	3169	2997	3026	3424	3729	3793	3785	4347	3566
Speaker 12F	4061	4105	3999	4022	3934	4129	4019	3950	3781	3882	3842
Speaker 13F	4102	3911	3671	3346	3403	3722	3657	3705	3911	3954	3796
Speaker 14F	4053	3134	3110	3001	3452	3770	3747	3619	3809	3865	3871
Mean	3939	3561	3372	3266	3402	3708	3717	3710	3764	3942	3609
St. Deviation	417	503	448	449	342	276	209	175	138	250	375

Table 7A: F4 of Vowels in Citation Form

Vowels/F4	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Spch	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 9F	3102	3050	3251	2942	3128	2972	3212	3381	3458	3394	3147
Speaker 11F	3991	3609	2730	3123	3305	3573	3768	3952	3938	4255	3482
Speaker 12F	4026	4076	3860	3701	3944	3983	3738	3845	3786	3887	3822
Speaker 13F	4081	2958	3046	3446	3462	3449	3214	3847	3862	4122	3646
Speaker 14F	3631	3380	3267	3229	3396	3431	3522	3762	3872	3527	3480
Mean	3766	3414	3230	3288	3447	3481	3490	3757	3783	3837	3515
St. Deviation	411	452	413	294	304	361	270	220	189	371	249

Table 7B: F3 of Vowels in Running Speech

Vowels/F4	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	3939	3561	3372	3266	3402	3708	3717	3710	3764	3942	3609
Running Speech	3766	3414	3230	3288	3447	3481	3490	3757	3783	3837	3515
Difference	173	147	142	22	45	236	227	47	19	105	94

Table 6C: Summary of F4 Differences

F4 measurements in female speech for CF and RS are respectively 3635 Hz and 3546 Hz. In male speech we have 3518 Hz and 3638 Hz (see Appendix 6). The differences are respectively 89 Hz and 92 Hz. Both are so far below the JND that we conclude here too that speech style variations do not affect the F4 of vowels.

4.0 Speech Style Variation and Intensity

Intensity correlates with the volume of the source of the sound/noise. Everything being equal, the larger the area of the emitting source, the greater the intensity. This explains why, generally speaking, sounds produced by males are louder than the ones produced by females. French and Steinberg (1947:100) provide the following statement about intensity which can also be applied to speech styles:

> Conversation at the rate of 200 words a minute, corresponding to about four syllables and ten speech sounds per second, is not unusual. During the brief period that a sound lasts, the intensity builds up rapidly, remains comparatively constant for a while, then decays rapidly.

The JND for the intensity domain is taken to be ≥ 3 dB by many reputable standardization organizations such as the National Institute for Occupational Safety and Health (NIOSH). However, it should be noted that this JND is only for speech intelligibility. Many sound level meters come pre-calibrated for either 3 dB or 5 dB settings. The latter is used for testing hearing acuity or for calculating noise pollution and environmental hazards. It is the JND used by the Occupational Safety and Health Administration (OSHA). Again, we want to emphasize here that we are dealing **only** with speech intelligibility, not hearing acuity. Consequently, the JND is stated as follows:

JND in the Intensity Domain

> Vowels in CF are auditorily distinct from vowels in RS, and vice versa, if and only if the intensity difference between them is ≥ 3 dB.

The intensity analysis in Table 8A through 8C yielded the following measurements:

Vowels/Ints	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	61	64	62	63	62	65	63	62	62	60	64
Speaker 2F	58	60	60	63	59	59	61	58	58	59	58
Speaker 3F	69	70	69	69	69	66	71	67	69	69	70
Speaker 4F	50	57	55	55	52	54	53	54	56	54	51
Speaker 5F	57	58	56	58	61	61	61	59	56	58	55
Speaker 6F	48	42	41	41	44	44	43	41	39	42	42
Speaker 7F	44	43	45	45	47	49	46	43	40	34	42
Speaker 8F	37	41	39	39	30	37	41	36	39	37	35
Speaker 9F	66	71	72	71	74	73	74	72	73	70	74
Speaker 10F	70	73	71	71	64	69	68	70	69	71	67
Speaker 15F	50	52	51	51	51	55	54	54	53	53	51
Speaker 16F	51	55	53	52	51	47	50	51	53	54	49
Speaker 17F	50	53	53	53	50	50	53	52	53	52	51
Mean	54	56	55	56	54	56	56	55	55	54	54
St. Deviation	9	10	10	11	10	10	11	11	11	11	11

Table 8A: Intensity of Vowels in Citation Form

Vowels/Ints	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Spch	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1F	65	64	61	67	68	66	61	63	67	60	65
Speaker 2F	52	55	51	56	53	55	54	50	56	55	54
Speaker 3F	68	65	63	68	65	64	61	65	66	65	66
Speaker 4F	49	53	49	57	55	51	53	54	57	53	54
Speaker 5F	56	57	58	61	62	59	59	58	60	57	57
Speaker 6F	43	44	41	48	46	47	46	44	46	45	44
Speaker 7F	41	37	42	47	49	51	35	39	44	42	46
Speaker 8F	45	43	41	45	41	44	41	43	43	44	42
Speaker 9F	68	72	72	75	74	75	70	74	74	72	75
Speaker 10F	68	69	66	72	72	73	68	67	72	72	73
Speaker 15F	53	51	52	54	54	52	51	54	53	54	52
Speaker 16F	54	55	52	57	56	55	54	54	56	54	53
Speaker 17F	56	55	55	55	52	52	55	56	57	54	53
Mean	55	55	54	58	57	57	54	55	57	55	56
St. Deviation	9	10	9	9	10	9	9	10	10	9	10

Table 8B: Intensity of Vowels in Running Speech

Vowels/Ints	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	54	56	55	56	54	56	56	55	55	54	54
Running Speech	55	55	54	58	57	57	54	55	57	55	56
Difference	1	1	1	2	3	1	2	0	2	1	2

Table 8C: Summary of Intensity Differences

The data provides us with mixed results. The intensity measurements of vowels in CF and RS are identical in female speech. They are both 55 dB. However, in male speech (see Appendix 7), the intensity of vowels in CF (64 dB) is 5dB louder than the intensity in RS (59 dB). This finding may have something to do with the anatomical configurations of male larynxes, or a variety of other issues, including but not limited to, room acoustics, microphone specifications, distance from microphone. Regardless, since intensity (loudness) is not phonemic in any human language, even if it were found to correlate with speech style, it would have no impact on intelligibility. The only important requirement is for the speaker to talk loudly enough for the microphones in the AI-enabled devices to pick up the speech signals being emitted.

5.0 Acoustic Vowel Spaces and Speech Style

Up until now, the main quest has been on the correlation between speech style and intelligibility. Now, we turn our attention to the correlation between speech style and acoustic vowel space. Do acoustic vowels spaces vary as a function of speech style? In answering this question, reference must be made to articulatory phonetic terms such as overshooting and undershooting. Both of these terms suggest that vowels have ideal articulatory targets (Thomas 2011:174, 277, 293). The targets are reached or not reached depending on the speech style. It is generally assumed that in CF (hypospeech, i.e., speaking slowly) vowels reach or even surpass their articulatory targets (overshooting), whereas in RS (hyperspeech, i.e., speaking fast) vowels do not reach their articulatory targets (undershooting) as noted in Baken and Orlikoff (2000:265). Consequently, in RS, vowels would tend to be centralized (Gahl et al. 2012:11). Do our data verify these claims? The acoustic vowel spaces in Figures 3 and 4 provide some answers.

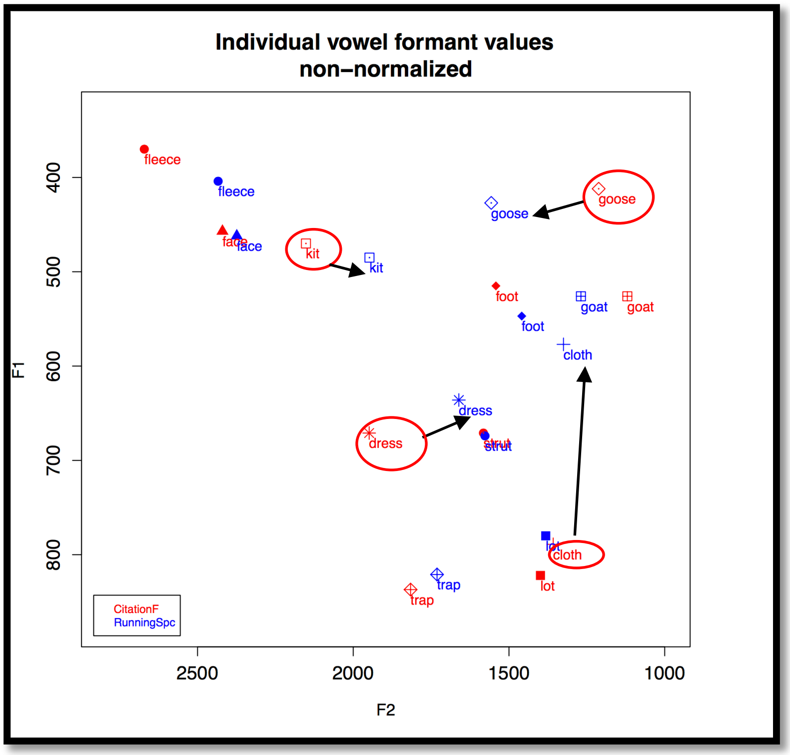

Figure 3: Acoustic Vowel Space for Females (Color Online)[9]

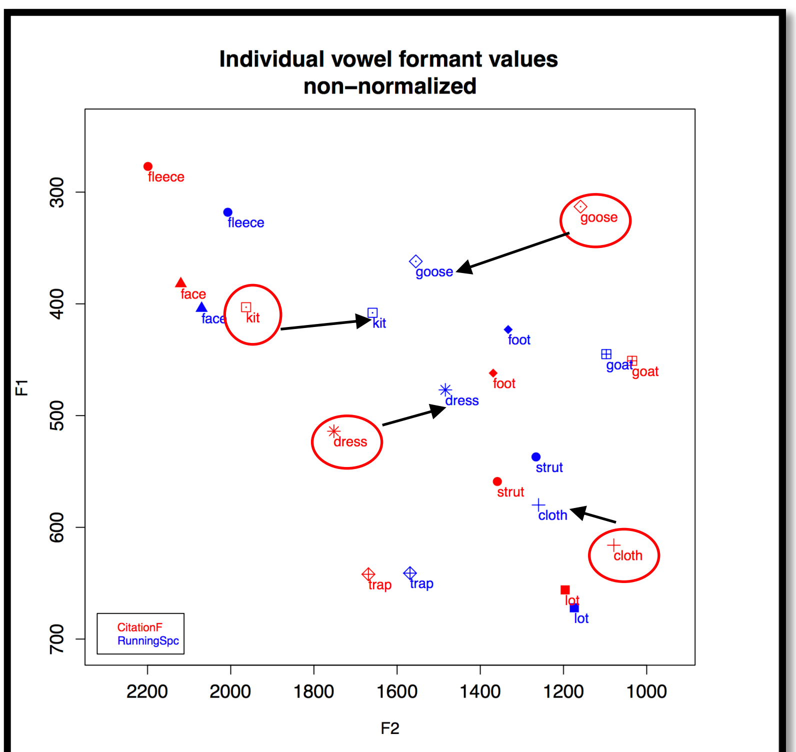

Figure 4: Acoustic Vowel Space for Males (Color Online)

[9] All acoustic vowel spaces are created through Norm (Thomas and Kendall 2007).

Indeed, in both female and male speech, vowels in RS display a proclivity towards centralization (undershooting), whereas vowels in CF are less likely to centralize. However, exceptions are easily found. Yaeger (1975:9) has a good explanation for these exceptions:

> In our analyses we have found quite generally that units not in a state of flux within a dialect are pronounced the same whether the token is measured from spontaneous or from controlled style, it falls within the same phonetic space.

On pages 14-15, she elaborates this point further:

> Thus far, I have illustrated the general principle that if a variable is not in a state of flux, the speaker will probably not vary his pronunciation of the unit in different styles; if on the other hand the vowel is in movement, an analysis of a range of styles will reveal that, at the very least, oscillation will take place at predictable points.

We see that in both female and male speech, the following vowels vary with regard to speech style: [ɪ, ɛ, u, ɔ]. Koffi (2014) and Koffi (2016) have documented that these vowels are in a state of flux in Central MN English. The variability of [ɔ] is remarkable not only because it is presently in active merger with [ɑ] but also because of its rise in RS versus CF. Its F1 in the former is 577 Hz versus 789 Hz in the latter. The difference of 212 Hz in F1 is perceptually significant. The rise does not amount to much in male speech where the difference between the two speech styles is only 36 Hz, that is, 580 Hz in RS versus 616 Hz in CF. The upward trajectory of [ɔ] has been noted by Labov et al. (2006:261) and by Chung (2020:542) for New York City and New Orleans. Women are at the forefront of this change in Central Minnesota English.

6.0 Summary and Implication for AI

From the preceding analyses, we conclude that the only correlates that varies as a function of speech style are duration, F0, and intensity (to some extent). Variations attributed to these correlates are inconsequential for intelligibility because neither of them can cause words to have different meanings in English. However, they are singled out additional consideration because humans are more acutely aware of them than the other correlates. In every day conversations with children or non-native speakers, they change their tempo, increase their pitch, and talk louder, all in an effort to maximize intelligibility. Because consumers do so almost instinctively, they may feel the need to do the same when using their AI-enabled devices. However, this is not necessary. AI-enabled devices can perceive speech signals at a normal speed, i.e., up to 200 wpm. Beyond this JND, intelligibility is likely to be problematic. In other words, unless one is speaking like an auctioneer (between 250 to 400 wpm), one does not need to change one's speech style at all when talking to Siri, Alexa, or Google Voice. More importantly, speakers need not talk louder than usual. So long as the volume of their voices is within 40 (quiet whisper) to 60 dB (normal conversation level, Fletcher 1953:77,100), the devices will understand them perfectly because the microphones in these devices are designed and manufactured to pick up speech signals below or beyond this intensity range. As for F1 and F2 formant frequencies, our study confirms what Ladefoged et al.(1976) and others uncovered more than 40 years ago, that is, variations in speech style have no corresponding impact on vowel formants. We see this not only for the two aforementioned formants, but also for F3 and F4, which previous studies did not investigate. In light of the present findings, it can be concluded that people can speak normally to their AI-enabled devices the same way they interact with human interlocutors.

ABOUT THE AUTHORS

Ettien Koffi, Ph.D. linguistics, teaches at Saint Cloud State University, MN. He is the author of four books and author/co-author of several dozen articles on acoustic phonetics, phonology, language planning and policy, emergent orthographies, syntax, and translation. His acoustic phonetic research is synergetic, encompassing L2 acoustic phonetics of English (Speech Intelligibility from the perspectives of the Critical Band Theory), sociophonetics of Central Minnesota English, general acoustic phonetics of Anyi (a West African language), acoustic phonetic feature extraction for application in Automatic Speech Recognition (ASR) and Text-to-Speech (TTS), and voice biometrics for speaker verification. He can be reached at enkoffi@stcloudstate.edu.

Jessica Krause is a graduate student in Audiology at the University of Minnesota, Twin Cities. She studied Communication Sciences and Disorders at Saint Cloud State University. She can be reached at krause.jesssica.jmk@gmail.com.

Reference

Baken, R. J. and Robert F. Orlikoff. 2000. *Clinical Measurement of Speech and Voice*. Second edition. San Diego, CA. Singular Publishing Group-Thomson Learning.
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/
Bradley Evan D. (2018). A Comparison of the Acoustic Vowel Spaces of Speech and Song. *Linguistic Research* 35 (2): 381-394.
Chung, Hyunju. 2020. Vowel Acoustic Characteristics of Southern American English Vowel Variation in Louisiana. *Journal of the Acoustical Society of America* 147 (1):541-553.
Crystal Thomas H and Arthur S. House. 1987. Segmental Durations in Connected Speech Signals: Current Results. *JASA* 83 (4):1553-1573.
Delattre, Pierre. 1951. The Physiological Interpretation of Sound Spectrograms *Proceedings of MLA*, Volume 66 (5):864-875.
Everest, Alton F. and Ken C. Pohlmann. 2015. *Master Handbook of Acoustics.* 6th edition. New York: McGraw Hill Education.
Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton & Co.-S-Gravenhage.
Fastl, Hugo and Eberhard Zwicker. 2007. *Psycho-Acoustics: Facts and Models*. 3rd edition. Munich, Germany: Springer
Fletcher, Harvey. 1953. Speech and Hearing in Communication. New York: D. Van Nostrand Company, Inc.
Fogerty, Daniel and Larry E. Humes. 2012. The Role of Vowel and Consonant Fundamental Frequency, Envelop, and Temporal Fine Structure Cues to the Intelligibility of Words and Sentences. Journal of the *Acoustical Society of America* 131 (2): 1490-1501.
French N. R and J.C. Steinberg. 1947. Factors Governing the Intelligibility of Speech Sounds. Journal of the *Acoustical Society of America* 19 (1): 90-119.
Fromkin, V. R. Rodman, and N. Hyams. 2014. An Introduction to Language. 10th edition. Boston, MA: Cengage Learning.

Gahl, Susanne, Yao Yao, and Keith Johnson. 2012. Why Reduce? Phonological Neighborhood Density and Phonetic Reduction in Spontaneous Speech. *Journal of Memory and Language* (Article in Print): 1-18.

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

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

Kewley-Port, Diane, T. Zachary Burkle, and Jae Hee Lee. 2007. Contribution of Consonant versus Vowel Information to Sentence Intelligibility for Young Normal-Hearing and Elderly Hearing-Impaired Listeners. *Journal of the Acoustical Society of America* 122 (4):2365-2375.

Koffi, Ettien. 2016. The Lowering of High Lax Vowels in Central Minnesota English: Does it Occur in other Dialects? *Linguistic Portfolios* 5: 2-14.

Koffi, Ettien. 2014. The Acoustic Vowel Space of Central Minnesota English in Light of the Northern Cities Vowel Shift. *Linguistic Portfolios* 3: 2-20.

Labov, William; Sharon Ash, and Charles Boberg. 2006. The Atlas of North American English: Phonetics, Phonology, and Sound Change.

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

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.

Ladefoged, Peter. 2006. *A Course in Phonetics.* Fifth Edition. Boston, MA: Thomson-Wadsworth.

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

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.

Ladefoged, Peter and D.E. Broadbent. 1957. Information Conveyed by Vowels. *Journal of the Acoustical Society of America* 29 (1):98-104.

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.

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.

Pope, J. 1998. Analyzers. *Handbook of Acoustics*, pp. 1341-1354. Ed by. Malcom J. Crocker. New York: A Wiley-Interscience Publication, John Wiley & Sons, Inc.

Rabiner, Lawrence and Biing-Hwang Juang. 1993. *Fundamentals of Speech Recognition*. Englewood Cliffs, New Jersey: Prentice Hall.

Repp, Bruno. 1987. Integration and Segregation in Speech Perception. *Haskins Laboratories Report on Speech Research SR 89/90*, pp. 1-32.

Scharf, Bertram. 1961. Complex Sounds and Critical Bands. *Psychological Bulletin* 58 (3):205-215.

Moon, Seung-Jae and Bjorn Lindblom. 1994. Interactions between Duration, Context, and Speaking Style in English Stressed Vowels. *Journal of the Acoustical Society of America* 95 (1):40-55.

Stevens, Kenneth. 2000. *Acoustic Phonetics*. MIT Press.
Thomas, Erik R. 2011. *Sociophonetics: An Introduction*. New York: Palgrave Macmillan.
Thomas, Erik R. and Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite. [Online Resource: http://ncslaap.lib.ncsu.edu/tools/norm/]
Ryant, Neville and Mark Liberman. 2016. Automatic Analysis of Phonetic Speech Style Dimensions. *Interspeech* 2016: 77-81.
Weinberger, Steven. (2019). *Speech Accent Archive*. George Mason University. Retrieved from http://accent.gmu.edu
Yaeger, Malcah. 1975. Speaking Style: Some Phonetic Realizations and their Significance. *Pennsylvania Working Papers on Linguistic Change* 1:1-61.

Native Speaker Males

Appendix 1: Duration

Vowels/Duration	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 2M	213	207	252	189	218	252	259	245	197	197	149
Speaker 3M	250	175	264	142	222	175	288	244	135	123	115
Mean	231	191	258	165	220	213	273	244	166	160	132
St. Deviation	26	22	8	33	2	54	20	0.7	43	52	24

Appendix 1A: Duration of Vowels in Citation Form

Vowels/Duration	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 2M	153	100	111	120	176	189	95	103	81	178	105
Speaker 3M	178	98	114	80	146	211	130	45	81	166	105
Mean	165	99	112	100	161	200	112	74	81	172	105
St. Deviation	17	1	2	28	21	15	24	41	0	8	0

Appendix 1B: Duration of Vowels in Running Speech

Vowels/Duration	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	231	191	258	165	220	213	273	244	166	160	132
Running Speech	165	99	112	100	161	200	112	74	81	172	105
Difference	66	92	146	65	59	13	161	170	85	12	27

Appendix 1C: Summary of Duration Differences

Appendix 2: F0/Pitch

Vowels/F0	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	123	113	111	109	122	108	113	104	125	124	110
Speaker 2M	143	126	123	120	121	119	116	117	128	132	112
Speaker 3M	135	132	130	131	129	128	130	131	139	144	133
Speaker 4M	90	93	90	92	88	85	88	83	87	154	94
Speaker 5M	84	82	80	79	79	78	78	80	82	82	85
Mean	115	109	106	106	107	103	105	103	112	127	106
St. Deviation	26	21	21	20	22	21	21	21	25	27	18

Appendix 2A: F0 of Vowels in Citation Form

Vowels/F0	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	115	127	106	117	104	95	101	134	120	108	110
Speaker 2M	110	109	101	113	100	99	108	110	131	117	120
Speaker 3M	145	137	147	127	127	119	121	191	139	133	128
Speaker 4M	98	84	92	108	94	93	84	116	102	100	103
Speaker 5M	82	81	87	82	81	77	74	83	90	82	87
Mean	110	107	106	109	101	96	97	126	116	108	109
St. Deviation	23	25	23	16	16	15	18	40	20	19	15

Appendix 2B: F0 of Vowels in Running Speech

Vowels/F0	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	115	109	106	106	107	103	105	103	112	127	106
Running Speech	110	107	106	109	101	96	97	126	116	108	109
Difference	5	2	0	3	6	7	8	23	4	19	3

Appendix 1C: Summary of F0 Differences

Appendix 3: F1

Vowels/F1	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	281	445	381	615	786	688	699	462	421	295	585
Speaker 2M	296	378	399	472	660	720	719	468	676	341	707
Speaker 3M	305	438	420	552	653	654	540	499	433	339	551
Speaker 4M	248	382	359	474	592	607	545	430	385	322	475
Speaker 5M	255	374	351	457	523	611	580	396	397	272	479
Mean	277	403	382	514	642	656	616	451	462	313	559
St. Deviation	24	34	28	67	97	48	86	39	120	29	94

Appendix 3A: F1 of Vowels in Citation Form

Vowels/F1	fleece	kit	face	dress	trap	lot	cloth	goat	foot	goose	strut
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	317	437	396	564	757	721	529	461	518	384	613
Speaker 2M	337	407	492	529	747	818	812	606	443	439	591
Speaker 3M	332	413	390	464	666	585	505	391	392	362	527
Speaker 4M	300	391	379	396	522	565	481	371	366	302	471
Speaker 5M	308	392	363	433	515	671	574	399	397	324	483
Mean	318	408	404	477	641	672	580	445	423	362	537
St. Deviation	15	18	50	68	117	103	134	95	59	53	63

Appendix 3B: F1 of Vowels in Running Speech

Vowels/F1	fleece	kit	face	dress	trap	lot	cloth	goat	foot	goose	strut
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	277	403	382	514	642	656	616	451	462	313	559
Running Speech	318	408	404	477	641	672	580	445	423	362	537
Difference	41	5	22	37	1	16	36	6	39	49	22

Appendix 3C: Summary of F1 Differences

Appendix 4: F2

Vowels/F2	fleece	kit	face	dress	trap	lot	cloth	goat	foot	goose	strut
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	2321	1952	2212	1847	1735	1195	1102	1011	1321	1282	1281
Speaker 2M	2278	1998	2136	1801	1868	1290	1085	1026	1669	1115	1502
Speaker 3M	2337	1860	2057	1723	1579	1283	1109	1163	1368	1073	1511
Speaker 4M	1747	2103	2003	1717	1600	1079	970	943	1131	1145	1193
Speaker 5M	2313	1902	2196	1676	1566	1133	1132	1033	1356	1183	1310
Mean	2199	1963	2120	1752	1669	1196	1079	1035	1369	1159	1359
St. Deviation	253	93	89	69	129	92	63	79	193	79	141

Appendix 4A: F2 of Vowels in Citation Form

Vowels/F2	fleece	kit	face	dress	trap	lot	cloth	goat	foot	goose	strut
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	2124	1755	2140	1465	1560	1220	1171	940	1348	1355	1243
Speaker 2M	2240	1784	2122	1636	1710	1388	1776	1572	1534	1904	1409
Speaker 3M	2051	1600	2024	1480	1512	1065	871	979	1287	1230	1251
Speaker 4M	1492	1505	1984	1285	1518	1011	1135	954	1228	1851	1128
Speaker 5M	2128	1651	2080	1557	1546	1189	1350	1043	1269	1438	1303
Mean	2007	1659	2070	1484	1569	1174	1260	1097	1333	1555	1266
St. Deviation	295	114	65	130	81	147	335	268	120	303	101

Appendix 4B: F2 of Vowels in Running Speech

Vowels/F2	fleece	kit	face	dress	trap	lot	cloth	goat	foot	goose	strut
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	2199	1963	2120	1752	1669	1196	1079	1035	1369	1159	1359
Running Speech	2007	1659	2070	1484	1569	1174	1260	1097	1333	1555	1266
Difference	192	304	50	268	100	22	181	62	36	396	93

Appendix 4C: Summary of F0 Differences

Appendix 5: F3

Vowels/F3	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	3203	2973	3095	2883	2578	2364	2419	2468	2539	2515	2608
Speaker 2M	2998	2579	2544	2553	2670	2424	2541	2426	2663	2369	2596
Speaker 3M	2858	2617	2676	2511	2379	2527	2515	2412	2434	2421	2440
Speaker 4M	2736	2788	2775	2752	2730	2573	2707	2537	2591	2592	2670
Speaker 5M	2828	2594	2643	2635	2556	2545	2445	2516	2409	2337	2460
Mean	2924	2710	2746	2666	2582	2486	2525	2471	2527	2446	2554
St. Deviation	181	169	211	151	133	88	113	54	106	105	99

Appendix 5A: F3 of Vowels in Citation Form

Vowels/F3	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	2889	2482	2838	2526	2366	2287	2428	2755	2606	2571	2401
Speaker 2M	2835	2067	2769	2699	2603	2655	2955	2940	2715	2786	2529
Speaker 3M	2603	2450	2528	2568	2207	2205	2191	2543	2483	2607	2291
Speaker 4M	2651	2779	*3068*	2839	2834	2711	2753	*2668*	1693	3128	2831
Speaker 5M	2577	2539	2587	2507	2449	2471	2670	2781	2523	2483	2529
Mean	2711	2463	2758	2627	2491	2465	2599	2737	2404	2715	2516
St. Deviation	141	256	214	139	238	221	296	146	407	255	202

Appendix 5B: F3 of Vowels in Running Speech

Vowels/F3	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	2924	2710	2746	2666	2582	2486	2525	2471	2527	2446	2554
Running Speech	2711	2463	2758	2627	2491	2465	2599	2737	2404	2715	2516
Difference	213	247	12	39	91	21	74	266	123	269	38

Appendix 5C: Summary of F3 Differences

Appendix 6: F4

Vowels/F4	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	3574	3360	3414	3462	3329	3039	3035	3353	3416	3369	3342
Speaker 2M	3714	3735	3752	3722	3653	3594	3403	3633	3744	3634	3610
Speaker 3M	3476	3489	3642	3622	3843	3657	3106	3120	3158	3042	3720
Speaker 4M	3458	3632	3827	3694	3605	3572	3670	3575	3441	3577	3811
Speaker 5M	3607	3569	3703	3732	3583	3509	3492	3409	3413	3338	3547
Mean	3565	3557	3667	3646	3602	3474	3341	3418	3434	3392	3606
St. Deviation	104	142	157	111	184	248	266	202	208	233	179

Appendix 6A: F4 of Vowels in Citation Form

Vowels/F4	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 1M	3513	3386	3366	3311	3222	3069	3392	3292	3452	3532	3276
Speaker 2M	3973	3898	3764	3801	3861	3727	4084	4186	3717	3827	3871
Speaker 3M	3542	3646	3423	3428	3853	4023	3438	3499	3310	3492	3612
Speaker 4M	3581	3889	3909	3776	3956	3796	3764	3569	3731	4041	3792
Speaker 5M	3716	3618	3742	3685	3720	3427	3808	3662	3480	3422	3838
Mean	3665	3687	3640	3600	3722	3608	3697	3641	3538	3662	3677
St. Deviation	188	213	234	218	292	369	285	333	181	261	245

Appendix 6B: F4 of Vowels in Running Speech

Vowels/F4	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	3565	3557	3667	3646	3602	3474	3341	3418	3434	3392	3606
Running Speech	3665	3687	3640	3600	3722	3608	3697	3641	3538	3662	3677
Difference	100	130	27	46	120	134	356	223	104	270	71

Appendix 6C: Summary of F4 Differences

Appendix 7: Intensity

Vowels/Intensity	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Citation Form	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 2M	57	57	55	51	53	54	55	52	58	55	51
Speaker 3M	74	74	74	76	75	77	75	76	77	77	77
Mean	65	65	64	63	64	65	65	64	67	66	64
St. Deviation	12	12	13	17	15	16	14	16	13	15	18

Appendix 7A: Intensity of Vowels in Citation Form

Vowels/ Intensity	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Running Speech	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Speaker 2M	52	48	49	47	47	45	42	45	54	48	50
Speaker 3M	71	75	75	74	73	74	69	69	71	70	76
Mean	61	61	62	60	60	59	55	57	62	59	63
St. Deviation	13	19	18	19	18	20	19	16	12	15	18

Appendix 7B: Intensity of Vowels in Running Speech

Vowels/ Intensity	**fleece**	**kit**	**face**	**dress**	**trap**	**lot**	**cloth**	**goat**	**foot**	**goose**	**strut**
Averages	[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Citation Form	65	65	64	63	64	65	65	64	67	66	64
Running Speech	61	61	62	60	60	59	55	57	62	59	63
Difference	4	4	2	3	4	6	11	8	5	7	1

Appendix 7C: Summary of Intensity Differences