[Linguistic Portfolios](https://repository.stcloudstate.edu/stcloud_ling)

Volume 10 Article 6

2021

Acoustic Phonetic Evidence of Masking Between [ʌ] and [ɔ] in Central Minnesota English

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

Connor Hanzsek-Brill
St. Cloud State University, cdhanzsekbrill@go.stcloudstate.edu

Recommended Citation

Koffi, Ettien and Hanzsek-Brill, Connor (2021) "Acoustic Phonetic Evidence of Masking Between [ʌ] and [ɔ] in Central Minnesota English," *Linguistic Portfolios*: Vol. 10 , Article 6.
Available at: https://repository.stcloudstate.edu/stcloud_ling/vol10/iss1/6

ACOUSTIC PHONETIC EVIDENCE OF MASKING BETWEEN [ʌ] AND [ɔ] IN CENTRAL MINNESOTA ENGLISH

ETTIEN KOFFI AND CONNOR HANZSEK-BRILL[1]

ABSTRACT

Since the Great Vowel Shift began in the 1400s, English vowels have been in a state of flux (Fromkin et al. 2017:331). Labov et al. (2006) have documented two macrolevel shifts in North American English: The Northern Cities Shift and The Southern Shift. At a microlevel, vowel shifts are happening in various regions of the US, as amply described in Wolfram and Ward (2006). However, vowel shifts in Minnesota English are poorly documented, except for stereotypical portrayals in popular culture. A shift involving [ʌ] and [ɔ] is quietly underway. It drew our attention only because of a misunderstanding between a speaker and dozens of listeners. Two separate experiments are conducted to find out more about what is going on with [ʌ] and [ɔ]. For this, we recruited 27 participants and extracted F0, F1, F2, F3, F4, intensity, and duration values, for a total of 436 measured tokens. The main finding is that many speakers in Central Minnesota no longer differentiate between [ʌ] and [ɔ] when they occur before nasal consonants.

Keywords: Masking Analysis, Vowel Masking, Confusion Analysis, Vowel Intelligibility, Vowel Shift, Minnesota English.

1.0 Introduction

The catalyst for this paper is a misunderstanding that occurred between native speakers of Central Minnesota English (CMNE). It happened when my daughter was in middle school on the Swim and Dive team. Booster Club members sold <mums>[2] for a fundraiser. On the last meet of the season, the Booster Club president made the following announcement to the team: *"Don't forget to pick up your mums. They are waiting outside."* The girls were confused and puzzled by this announcement because they were not of driving age yet. Furthermore, why would their <moms> be waiting outside in the cold instead of coming inside? Seeing the confused looks on the girls' faces, the secretary of the Booster Club corrected the announcement with the following, "Your 'mums' <m.u.m.s> not your <m.o.m.s>." The spelling helped to eliminate the confusion. On the drive home, I asked my daughter and her three friends, all who are from Central Minnesota, why they were confused. In the conversation that followed, I asked them if they pronounced <mums> vs. <moms> and <sun> vs. <son> the same or differently. A lively debate ensued between the girls. Two said that they pronounced them the same, while two others opined that they said them differently. Since then, I have asked dozens of Minnesotans the same questions and have received the same mixed answers. This acoustic phonetic study is undertaken to clarify the issue of how Minnesotans produce [ʌ] and [ɔ] before nasal segments.

[1]**Authorship responsibilities:** Author 1 assigned this topic to Author 2 who was enrolled in his acoustic phonetics course. Author 2 collected data from 20 participants for Experiment 1. Author 1 also collected data from seven participants for Experiment 2. Author 2 is recognized as such because of the data he collected and the measurements he did for Experiment 1. Author 1 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. To the extent that the measurements provided by Author 2 are accurate, they both share equally the rights, privileges, and responsibilities of this publication.

[2] By linguistic convention, the angle brackets <…> represent the orthographic form of words.

2.0 Research Design and Methodology

The investigation involves two separate experiments. The first was conducted mostly by the second author (Author 2). He needed Independent Study credits to complete his BA in linguistics. Since he had taken the first author's (Author 1) acoustic phonetics course, he asked Author 2 to investigate the issue of [ʌ] and [ɔ] for his capstone project. He recruited 20 participants (10 females and 10 males) from Central Minnesota. Author 2 secured the Institutional Review Board (IRB) approval for Experiment 1. The participants consented and were asked to read the following sentences as naturally as possible:

Sentence 1: *Angry fish book.*
Sentence 2: *Chewing food with closed mouths.*
Sentence 3: *That sun is bright.*
Sentence 4: *The mouse is running up the wall.*
Sentence 5: *Happy buffalos eat corn.*
Sentence 6: *Chewing food with closed moths.*
Sentence 7: *Angry fish buck.*
Sentence 8: *The mouse is running up the wool.*
Sentence 9: *That son is bright.*
Sentence 10: *Happy buffalos ate corn.*

Subsequently, Author 1 carried out Experiment 2. The 13 participants who took part in this experiment were enrolled in Author 1's *Introduction to Linguistics* course. The class was an eclectic mix of majors: Criminal Justice/Forensic Science, Communication Sciences and Disorders, Communication Arts and Literature, World Languages, Teaching English as a Second Language, and Linguistics. They were asked to record themselves reading the four sentences below as naturally as possible:

Sentence 1: *That son is bright*.
Sentence 2: *That sun is bright*.
Sentence 3: *Pick up your moms*.
Sentence 4: *Pick up your mums*.

The participants were explicitly told that Author 1 was conducting this research to find out whether or not CMNE produce [ʌ] and [ɔ] identically. Only the measurements from the participants who grew up in Central Minnesota are used in this paper. Collectively, the 27 participants in Experiments 1 and 2 produced 436 tokens of [ʌ] and [ɔ].

3.0 Masking and Speech Intelligibility

Speech intelligibility researchers working from the perspective of the Critical Band Theory (CBT) have uncovered important Just Noticeable Difference (JND) thresholds at which speech signals mask each other and infringe on intelligibility. Masking is defined as follows:

Acoustic Masking and Intelligibility (AMI)

Segments that are acoustically close may mask each other with only a minimal risk to intelligibility, unless their relative functional loads dictate otherwise (Koffi (2021:55).

The masking analysis involves the correlates and JNDs listed below:

JND in the F0 Domain

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the pitch difference between them is ≥ 1 Hz on the F0 frequency bandwidth. Otherwise, masking is likely.

JND in the F1 Domain

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the F1 difference between them is ≥ 60 Hz on the F1 frequency bandwidth. Otherwise, masking is likely.

JND in the F2 Domain

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the F2 difference between them is ≥ 200 Hz on the F2 frequency bandwidth. Otherwise, masking is likely.

JND in the F3 Domain

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the F3 difference between them is ≥ 400 Hz on the F3 frequency bandwidth. Otherwise, masking is likely.

JND in the F4 Domain

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the F4 difference between them is ≥ 600 Hz on the F4 frequency bandwidth. Otherwise, masking is likely.

JND in the Intensity Domain

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the intensity difference between them is ≥ 3 dB in the intensity domain. Otherwise, masking is likely.

JND in the Duration Domain-1

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the duration difference between them is ≥ 10 ms in the temporal domain for signals lasting less than 200 ms. Otherwise, masking is likely.

JND in the Duration Domain-2

Two speech signals **A** and **B** are auditorily distinct from each other if and only if the duration difference between them is ≥ 20 ms in the temporal domain for signals lasting less than 300 ms. Otherwise, masking is likely.

These JNDs are scattered all over the expansive psychoacoustic literature but Koffi (2021:38-40) has conveniently summarized them for ease of reference. The JNDs of F1 and F2 are the ingredients on which Labov et al.'s (2006) based their interpretation of vowels shifts in *Atlas of North American English*, Labov et al. (2013), and other publications. In using these JNDs, it is good to keep in mind that F1 plays a more prominent role in intelligibility than all other formants because it alone contains 80% of the acoustic energy found in vowels (Ladefoged and Johnson 2015:207). For this reason, F1 will be used in accordance with the four severity scales discussed in Koffi (2021:75) to understand how and why [ʌ] and [ɔ] mask each other.

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%	Mediocre intelligibility
4.	0 Hz – 20 Hz	Complete masking	75–100%	Poor intelligibility

Table 1: Severity Scale of Masking and Intelligibility

Furthermore, Catford (1987) provides relative functional load (RFL) calculations for [ʌ] and [ɔ] which helps to estimate intelligibility ratings. Since the RFL of [ʌ] vs. [ɔ] is 65%, it means that confusion between the two leads to mediocre intelligibility, as indicated in Table 1.

3.1 Review of the Literature about the Confusion between [ʌ] and [ɔ]

Peterson and Barney (1952) conducted a seminal research on the production and perception of American English vowels. They recruited 76 participants: 33 men, 38 women, and 15 children. They spoke General American English (GAE)[3]. One aspect of their research dealt with the results of a confusion study (see page 182). The findings reveal that the hearers confused [ʌ] with [ɔ] 1.23% of the time, and confused [ɔ] with [ʌ] 1.66% of the time. The two-way confusion rate yields 2.89%. Ladefoged and Disner (2012:43) refer to GAE as "a more old-fashioned dialect." Even so, the findings reveal that as far back as 1952, some 70 years ago, people confused [ʌ] and [ɔ]. Hillenbrand et al. (1995:3108) replicated Peterson and Barney's study to investigate vowels produced by Midwesterners. They recruited 139 participants from all over the Midwest, including two from Minnesota. However, 87% of their participants were from Michigan's lower peninsula. Their findings were that [ʌ] was confused with [ɔ] 1.8% of the time, and [ɔ] was confused with [ʌ] 3.8% of the time. The overall rate of confusion was 5.6%. Their study took place 26 years ago. Since, the rate of confusion between [ʌ] and [ɔ] is on the increase in the Midwest, it is reasonable to expect this to be the case in Minnesota. This is what we are trying to find out by conducting Experiments 1 and 2.

3.2 Masking of [ʌ] and [ɔ] in GAE

Where there is confusion, there is masking, and vice versa. In confusion studies, researchers elicit hearers' responses as to whether or not they auditorily perceive differences between two speech sounds. Masking studies, on the other hand, examine acoustic measurements obtained from talkers to see whether or not speech signals overlap in auditory space. The degree of overlapping helps gauge the severity of masking and confusion. In addition to providing confusion results, Peterson and Barney (1952) measured various acoustic correlates of GAE vowels. We focus only on the measurements of [ʌ] and [ɔ] to gauge levels of masking. The JNDs listed in 3.0 are the yardsticks used to make masking assessments. The relevant correlates are displayed in Table 2:

	F0 [ʌ]	F0 [ɔ]	F1 [ʌ]	F1 [ɔ]	F2 [ʌ]	F2 [ɔ]	F3 [ʌ]	F3 [ɔ]
Peterson and Barney (1952:183) M	130	129	640	570	1190	840	2390	2410
Peterson and Barney (1952:183) W	221	216	760	590	1400	920	2780	2710

Table 2: Measurements of the Strut Vowel [ʌ] and [ɔ] in GAE

No masking occurs on the F0 frequency bandwidth because the acoustic distance between [ʌ] and [ɔ] is 1 Hz in male speech. In female speech, it is 5 Hz. Masking does not take place on the F1 frequency bandwidth either because [ʌ] is distant from [ɔ] by 70 Hz in male speech, and by 170 Hz in female speech. There is also no masking on the F2 frequency bandwidth because [ʌ] and [ɔ] are separated by 350 Hz in male speech, and by 480 Hz in female

[3] Peterson and Barney (1952:177) used this phrase only for male speakers.

speech. The only correlate that shows masking is F3. The acoustic distance between [ʌ] and [ɔ] is 20 Hz in male speech and 70 Hz in female speech. Since these measurements are below the JND of 400 Hz that is required for optimal auditory perception on the F3 bandwidth, we conclude that [ʌ] and [ɔ] mask each other on this bandwidth. However, F3 is inconsequential for intelligibility because, as Ladefoged and Disner (2012:46) note, "[it] has very little function in distinguishing vowels." Since, there is no masking in F1 or F2, it is unclear as to why the participants confused [ʌ] with [ɔ] and [ɔ] with [ʌ] at a rate of 2.89%.

3.3 Masking of in [ʌ] and [ɔ] in the Midwest

Hillenbrand et al. (1995) replicated Peterson and Barney's study to investigate Midwest vowels. They provided F4 and duration measurements, which were missing from Peterson and Barney's paper. The six correlates that pertain to [ʌ] and [ɔ] from the study of Midwest vowels are displayed in Table 3A and 3B:

	F0 [ʌ]	F0 [ɔ]	F1 [ʌ]	F1 [ɔ]	F2 [ʌ]	F2 [ɔ]	F3 [ʌ]	F3 [ɔ]
Hillenbrand et al (1995:3103) M	133	121	623	652	1200	997	2550	2538
Hillenbrand et al (1995:3103) W	218	210	753	781	1426	1136	2933	2828

Table 3A: Measurements of [ʌ] and [ɔ] in Midwest English

	F4 [ʌ]	F4 [ɔ]	Duration [ʌ]	Duration [ɔ]
Hillenbrand et al (1995:3103) M	3557	3486	188	283
Hillenbrand et al (1995:3103) F	4092	3923	226	353

Table 3B: Measurements of [ʌ] and [ɔ] in Midwest English

Hillenbrand et al. (1995) reported a confusion rate of 5.6% of the time between [ʌ] and [ɔ]. We note in passing that the rate of confusion is almost double the one in GAE.

Masking does not occur on the F0 frequency bandwidth because the acoustic distances between [ʌ] and [ɔ] are respectively 12 Hz and 8 Hz in male and female speech. However, masking is evident between [ʌ] and [ɔ] on the F1 frequency bandwidth because the acoustic distances between the two vowels are respectively 29 Hz and 28 Hz for men and women. According to the scales in Table 1, this corresponds to moderate masking. The evidence of masking in F1 alone is sufficient to explain why Midwesterners confuse [ʌ] and [ɔ]. As noted in 3.0, F1 accounts for 80% of the acoustic energy in vowels. There is no masking on the F2 frequency bandwidth because the acoustic distances between the two vowels are respectively 203 Hz in male speech and 290 Hz in female speech. Since these values are higher than the JND of 200 Hz, we infer that the participants did not form a homogeneous dialect block. If their dialect were the same, the F2 measurement between [ʌ] and [ɔ] would have been less or equal to 200 Hz. Our inference is correct because we know that 119 of the 139 participants were from Michigan alone. The remaining 20 participants were from five states: Illinois, Wisconsin, Minnesota, Ohio, and Indiana, probably two participants per state (p. 3100). There is evidence of masking in F3 and F4. However, as already stated, these higher formants play only a marginal role, if any, on intelligibility. Masking does not occur in the duration domain because the temporal distance between [ʌ] and [ɔ] are 95 ms in male speech and 127 ms in female speech. This also means that speakers from the Midwest can use duration to encode or decode differences between [ʌ] and [ɔ].

4.0 Masking Analyses Based on Experiment 1

Experiment 1 is not a replication of the two previous studies. Yet, the same correlates found in Hillenbrand et al. (1995) were extracted from 10 males and 10 females. Additionally, intensity measurements were collected. Experiment 1 was run entirely by Author 2 who collected and measured all the data. Author 1 only checked to verify that nothing was amiss and was satisfied with the measurements.

4.1 Masking Analysis of F0 Based on Experiment 1

The anatomical correlate of F0 is vocal fold vibrations, which hearers perceive as pitch. Since English is an accent language (not a tone language), pitch plays no role in the intelligibility of <sun> vs. <son> even if masking were to occur.

Word	**sun**	**son**	**Word**	**sun**	**son**
Vowel	[ʌ̃]	[ɔ̃]	Vowel	[ʌ̃]	[ɔ̃]
Correlate	**F0**	**F0**	**Correlate**	**F0**	**F0**
Speaker 1F	230	232	Speaker 1M	137	135
Speaker 2F	203	201	Speaker 2M	146	126
Speaker 3F	258	237	Speaker 3M	98	104
Speaker 4F	240	234	Speaker 4M	155	173
Speaker 5F	193	182	Speaker 5M	130	143
Speaker 6F	192	192	Speaker 6M	130	130
Speaker 7F	207	203	Speaker 7M	130	121
Speaker 8F	247	223	Speaker 8M	155	211
Speaker 9F	217	195	Speaker 9M	162	175
Speaker 10F	237	239	Speaker 10M	122	119
Mean	**222**	**213**	**Mean**	**136**	**143**
St. dev.	23	21	St. dev.	18	32

Table 4: F0 Measurements of [ʌ] and [ɔ] in CMNE Experiment 1

The arithmetic means show that masking does not occur because the acoustic distances between [ʌ] and [ɔ] are greater than 1 Hz. In fact, they are 9 Hz in female speech and 7 Hz in male speech.

4.1 Masking Analysis of F1 Based on Experiment 1

F1 correlates with vowel height (i.e., mouth aperture). In the traditional classification of vowels, both [ʌ] and [ɔ] are treated as mid-low vowels (Ladefoged and Johnson 2015:228). This means that their F1 measurements are expected to be greater than 600 Hz. The arithmetic means in Table 5 show that these expectations are borne out:

Word	**sun**	**son**	**Word**	**sun**	**son**
Vowel	[ʌ̃]	[ɔ̃]	Vowel	[ʌ̃]	[ɔ̃]
Correlate	**F1**	**F1**	**Correlate**	**F1**	**F1**
Speaker 1F	711	787	Speaker 1M	701	647
Speaker 2F	782	656	Speaker 2M	605	729
Speaker 3F	730	842	Speaker 3M	567	571
Speaker 4F	759	746	Speaker 4M	594	636
Speaker 5F	709	567	Speaker 5M	676	553
Speaker 6F	749	687	Speaker 6M	589	566
Speaker 7F	684	718	Speaker 7M	483	674
Speaker 8F	701	684	Speaker 8M	625	635
Speaker 9F	803	732	Speaker 9M	677	694

Speaker 10F	694	683	Speaker 10M	528	567
Mean	**732**	**710**	**Mean**	**604**	**627**
St. dev.	39	74	St. dev.	68	61

Table 5: F1 Measurements of [ʌ] and [ɔ] in CMNE Experiment 1

Masking does occur between [ʌ] and [ɔ] because their acoustic distances are 22 Hz in female speech and 23 Hz in male speech. This corresponds to moderate masking on the severity scale in Table 1, like in the Midwest (see 3.3). Since F1 plays a greater role in intelligibility than all the other formants, the reduced distance between [ʌ] and [ɔ] ineluctably translates into greater incidences of masking. This is in fact the case because masking occurs in the pronunciation of 11 out of 20 (55%) participants. More will be said about this in Discussions in 6.0.

4.2 Masking Analysis of F2 Based on Experiment 1

F2 correlates anatomically with horizontal tongue movements. When F2 frequencies are higher than 2000 Hz, the vowel is classified as a front vowel. If the measurements are 1,800 to 1,400 Hz, the vowel is deemed as central. Measurements below 1,400 Hz are indicative of back vowels. According to these criteria, both [ʌ] and [ɔ] are central vowels in female speech but are back vowels in male speech.[4]

Word	**sun**	**son**	**Word**	**sun**	**son**
Vowel	[ʌ̃]	[ɔ̃]	Vowel	[ʌ̃]	[ɔ̃]
Correlate	**F2**	**F2**	**Correlate**	**F2**	**F2**
Speaker 1F	1526	1503	Speaker 1M	1269	1340
Speaker 2F	1545	1611	Speaker 2M	1431	1377
Speaker 3F	1379	1176	Speaker 3M	1116	1193
Speaker 4F	1614	1555	Speaker 4M	1439	1491
Speaker 5F	1650	1572	Speaker 5M	1392	1391
Speaker 6F	1633	1588	Speaker 6M	1637	1600
Speaker 7F	1393	1469	Speaker 7M	1332	1345
Speaker 8F	1655	1501	Speaker 8M	1319	1295
Speaker 9F	1530	1367	Speaker 9M	1270	1279
Speaker 10F	1443	1365	Speaker 10M	1376	1344
Mean	**1536**	**1470**	**Mean**	**1358**	**1365**
St. dev.	103	133	St. dev.	136	113

Table 6: F2 Measurements of [ʌ] and [ɔ] in CMNE Experiment 1

The more interesting finding is that the arithmetic mean of the two vowels shows masking for 19 out of 20 (95%) participants, except Speaker 3F. This a confirmation that the participants form a homogeneous dialect group. In female speech, the acoustic distance between [ʌ] and [ɔ] is 66 Hz. In male speech, it is only 7 Hz. The fact that masking occurs in both F1 and F2 in Experiment 1 is an indication that greater incidences of confusion between [ʌ] and [ɔ] are to be expected among CMNE speakers.

4.3 Masking Analysis of F3 Based on Experiment 1

The anatomical correlate of F3 is lip position and the lowering of the velum. Raised velum correlates almost automatically with unrounded lips, whereas the lowering of the velum almost automatically causes the lips to be rounded. F3 values that are ≥ 2,600 Hz correspond

[4] It is generally accepted in acoustic phonetics that females' formants are 20% higher than males' (Heller 2013:370). This means that if the JNDs were adjusted to female speech, the vowels would most likely qualify as back.

to unrounded lips, measurements below this JND indicate that the lips are lightly or strongly rounded. The measurements in Table 5 match up perfectly with the traditional articulatory classification of vowels. In most textbooks, including Fromkin et al. (2017:233), [ʌ] is classified as [-round], while [ɔ] is given the feature [+round].

Word	**sun**	**son**	**Word**	**sun**	**son**
Vowel	[ʌ̃]	[ɔ̃]	Vowel	[ʌ̃]	[ɔ̃]
Correlate	**F3**	**F3**	**Correlate**	**F3**	**F3**
Speaker 1F	2659	2640	Speaker 1M	2795	2838
Speaker 2F	2145	2474	Speaker 2M	2442	2375
Speaker 3F	2282	1742	Speaker 3M	2655	2374
Speaker 4F	2817	2827	Speaker 4M	2557	2512
Speaker 5F	2633	2455	Speaker 5M	2884	3053
Speaker 6F	2741	2598	Speaker 6M	2870	2419
Speaker 7F	3002	2943	Speaker 7M	2477	2604
Speaker 8F	2773	2867	Speaker 8M	2377	2340
Speaker 9F	2677	2361	Speaker 9M	2895	2901
Speaker 10F	2986	2562	Speaker 10M	2336	2513
Mean	**2671**	**2546**	**Mean**	**2628**	**2592**
St. dev.	273	340	St. dev.	219	251

Table 7: F3 Measurements of [ʌ] and [ɔ] in CMNE Experiment 1

The arithmetic means show that masking occurs between [ʌ] and [ɔ] because the acoustic distances between them are 125 Hz in female speech, and 36 Hz in male speech. Masking occurs for 17 of the 20 participants (85%), except for Speaker 3F, 10F, and 6M. Again, F3 plays only a marginal role, if any, in intelligibility. Therefore, masking on this bandwidth is inconsequential for the intelligibility of [ʌ] and [ɔ].

4.4 Masking Analysis of Intensity Based on Experiment 1

Intensity is an important acoustic correlate of speech because talkers and hearers rely on sonority differences between segments to encode and decode speech signals. However, since intensity is not phonemic in any human language, it cannot be relied on to assess masking. This is probably the reason why it was not included in previous studies on the intelligibility of vowels.

Word	**sun**	**son**	**Word**	**sun**	**son**
Vowel	[ʌ̃]	[ɔ̃]	Vowel	[ʌ̃]	[ɔ̃]
Correlate	**Int**	**Int**	**Correlate**	**Int**	**Int**
Speaker 1F	62	61	Speaker 1M	66	67
Speaker 2F	63	61	Speaker 2M	64	66
Speaker 3F	68	68	Speaker 3M	65	64
Speaker 4F	70	70	Speaker 4M	64	68
Speaker 5F	68	66	Speaker 5M	67	66
Speaker 6F	61	64	Speaker 6M	67	66
Speaker 7F	66	64	Speaker 7M	66	68
Speaker 8F	67	68	Speaker 8M	71	76
Speaker 9F	72	67	Speaker 9M	70	73
Speaker 10F	65	66	Speaker 10M	64	64
Mean	**66**	**65**	**Mean**	**66**	**67**
St. dev.	3	2	St. dev.	2	3

Table 8: Intensity Measurements of [ʌ] and [ɔ] in CMNE Experiment 1

Eighteen of the 20 participants (90%) produced [ʌ] and [ɔ] with masking intensity levels, except Speakers 9F and 9M. Since, the arithmetic mean shows a difference of only 1 dB for female and male speakers, we conclude that [ʌ] and [ɔ] mask each other in intensity.

4.5 Masking Analysis of Duration Based on Experiment 1

Duration plays an important role in the auditory discrimination between speech segments. It is phonemic in some languages, but not in English.

Word	**sun**	**son**	**Word**	**sun**	**son**
Vowel	[ʌ̃]	[ɔ̃]	Vowel	[ʌ̃]	[ɔ̃]
Correlate	**Dur**	**Dur**	**Correlate**	**Dur**	**Dur**
Speaker 1F	68	75	Speaker 1M	115	155
Speaker 2F	52	69	Speaker 2M	86	101
Speaker 3F	133	124	Speaker 3M	74	126
Speaker 4F	130	105	Speaker 4M	90	105
Speaker 5F	70	52	Speaker 5M	90	101
Speaker 6F	92	130	Speaker 6M	144	158
Speaker 7F	94	111	Speaker 7M	48	61
Speaker 8F	113	71	Speaker 8M	61	79
Speaker 9F	83	160	Speaker 9M	142	181
Speaker 10F	154	205	Speaker 10M	63	64
Mean	**98**	**110**	**Mean**	**91**	**113**
St. dev.	32	47	St. dev.	33	41

Table 9: Duration Measurements of [ʌ] and [ɔ] in CMNE in Experiment 1

The JND for perceiving one speech signal as lasting longer than another is 10 ms for signals that are less than 200 ms. The arithmetic means show that [ʌ] and [ɔ] do not mask each other in the temporal domain. The durational distances between them are respectively 12 ms in female speech, and 22 ms in male speech. Seventeen of the 20 participants (85%) relied on temporal cues to distinguish between [ʌ] and [ɔ]. By "rely on," we do not mean to imply that the participants did so consciously. Yet, in a situation where differentiation is required, they can rely on temporal cues to do so.

4.6 Interim Conclusion

To recap, for the participants in Experiment 1, the vowels [ʌ] and [ɔ] mask each other in F1, F2, F3, and intensity. They do not mask each other in F0 and duration. Since English is not a tone language, the lack of masking in pitch does not result in higher or lower rates of intelligibility. F3 and intensity also do not play any major role in the intelligibility of vowels. Therefore, the fact that these two correlates mask each other does not mean much. However, the fact that the two vowels mask each other in F1 is extremely important because, as we have said it numerous times before, F1 alone contains 80% of the acoustic energy in vowels. The masking in F2 is important because it confirms that participants belong to the same dialect community.

5.0 Masking Analyses Based on Experiment 2

The design of Experiment 2 is different from Experiment 1 in that, whereas the goal of the research was hidden from the participants of Experiment 1, it was stated plainly to the participants in Experiment 2. It was up to them to prove that they produced [ʌ] and [ɔ] identically or differently. It was Author 1's expectation that they would put on their "best pronunciation behavior" and differentiate between the two vowels. It did not matter to Author

1 that they would do so because his primary goal was to investigate pronunciation tendencies. The research hypothesis is that, if in spite of being told explicitly what the research was about, if the arithmetic mean shows masking in F1, this would prove that CMNE talkers cannot discriminate between [ʌ] and [ɔ] in production, even when they monitor their speech. An element of Experiment 2 that was not explored in Experiment 1 is the inclusion of the pair <mums> vs. <moms>, which is the misunderstanding that set this research project in motion.

5.1 Masking Analysis of F0 Based on Experiment 2

The arithmetic mean in Table 8 indicates that the vowels [ʌ] and [ɔ] do not mask each other in F0 because the acoustic distances between them are respectively 1 Hz for <sun> vs. <son>, and 20 Hz for <mums> vs. <moms>. The larger difference between <mums> vs. <moms> is of no consequence for intelligibility because English is not a tone language.

F0	**Sun**	**Son**	**Mums**	**Moms**
Speaker 1F	186	199	233	208
Speaker 2F	220	232	210	231
Speaker 3F	274	279	209	204
Speaker 4F	241	249	191	215
Speaker 5F	257	210	218	188
Speaker 6F	264	268	256	167
Speaker 7F	250	244	222	214
Mean	**232**	**231**	**219**	**199**
St. Dev.	39	36	19	23

Table 10: F0 Measurements of [ʌ] and [ɔ] in CMNE in Experiment 2

5.2 Masking Analysis of F1 Based on Experiment 2

The arithmetic mean shows clearly that [ʌ] and [ɔ] mask each other because the acoustic differences between <sun> vs. <son> is 49 Hz, while that between <mums> vs. <moms> is 44 Hz. Again, it is worth repeating that F1 is extremely important for intelligibility. Masking in Experiment 2 indicates that CMNE speakers produce [ʌ] and [ɔ] similarly even when they monitor their speech. Masking in Table 11 is slight, but it is masking nonetheless.

F1	**Sun**	**Son**	**Mums**	**Moms**
Speaker 1F	461	489	434	457
Speaker 2F	918	940	850	1136
Speaker 3F	795	722	635	466
Speaker 4F	712	786	650	810
Speaker 5F	811	703	851	728
Speaker 6F	662	578	652	708
Speaker 7F	970	772	630	703
Mean	**712**	**761**	**715**	**671**
St. Dev.	146	170	229	143

Table 11: F1 Measurements of [ʌ] and [ɔ] in CMNE in Experiment 2

We see that three participants (1F, 2F, 6F) out of seven (43%) produced the two vowels identically as far as masking is concerned. Speaker 1F confused [ʌ] and [ɔ] in both <sun> vs. <son> and <mums> vs. <mums>. She is the only one who consistently confused the two vowels. Speaker 2F and 6F displayed intraspeaker variability in that the former confused [ʌ] and [ɔ] in <sun> vs. <son>, but not in <mums> vs. <moms>, while the latter confused them in <mums> vs. <moms>, but not in <sun> vs. <son>.

5.3 Masking Analysis of F2 Based on Experiment 2

The arithmetic mean indicates that masking occurs between [ʌ] and [ɔ] in <sun> vs. <son> (40 Hz) and <mums> vs. <moms> (73 Hz). Masking here confirms, once again, that CMNE is a homogeneous dialect.

F2	Sun	Son	Mums	Moms
Speaker 1F	1731	1662	1449	1468
Speaker 2F	2139	2135	1981	1989
Speaker 3F	1901	1978	1755	1462
Speaker 4F	1754	1902	1760	1686
Speaker 5F	1963	1925	2002	1761
Speaker 6F	1763	2035	1706	1869
Speaker 7F	1900	1790	1734	1641
Mean	**1918**	**1878**	**1696**	**1769**
St. Dev.	1156	145	195	185

Table 12: F2 Measurements of [ʌ] and [ɔ] in CMNE in Experiment 2

5.4 Masking Analyses of F3 and F4 Based on Experiment 2

Since F3 and F4 have such a negligible impact on intelligibility, there is no harm in grouping them together in this section. We will deal with F3 first. Anatomically, it correlates with lip rounding, as noted in 4.3. The arithmetic mean in Table 13 indicates that [ʌ] and [ɔ] in <sun> vs. <son> and <mums> vs. <moms> mask each other because their acoustic distances are 55 Hz and 51 Hz respectively, well below the JND of 400 Hz.

F3	Sun	Son	Mums	Moms
Speaker 1F	2862	2832	2655	2581
Speaker 2F	3402	3446	3234	3549
Speaker 3F	3345	3147	3192	2721
Speaker 4F	2884	2784	2911	3068
Speaker 5F	3169	3187	3232	3063
Speaker 6F	3016	3006	3085	3018
Speaker 7F	3138	3027	2888	2840
Mean	**3061**	**3116**	**2977**	**3028**
St. Dev.	225	210	312	218

Table 13: F3 Measurements of [ʌ] and [ɔ] in CMNE in Experiment 2

F4 correlates anatomically with the size of the speaker's head (Ladefoged and Johnson 2017:222)[5] and/or with the size of the laryngeal cavity (Cao and Dellwo 2019:4).

F4	Sun	Son	Mums	Moms
Speaker 1F	4217	4352	4295	4248
Speaker 2F	4291	4501	4247	4343
Speaker 3F	4352	4426	4361	4129
Speaker 4F	4270	4521	4138	3983
Speaker 5F	4364	4685	4127	3970
Speaker 6F	4028	4142	4131	4057
Speaker 7F	4106	3891	3902	3962
Mean	**4359**	**4232**	**4098**	**4171**
St. Dev.	265	125	149	149

Table 14: F4 Measurements of [ʌ] and [ɔ] in CMNE in Experiment 2

[5] According to Ladefoged (2006:187) people with bigger heads tend to have slightly lower F4 values.

F4 differences between [ʌ] and [ɔ] in <sun> vs. <son> is 127 Hz, while that between <mums> vs. <moms> is 73 Hz. Since these values are well below the JND of 600 Hz, we conclude that masking occurs.

5.5 Masking Analysis of Intensity Based on Experiment 2

The sonority distances between [ʌ] and [ɔ] show that these two vowels mask each other in intensity. The difference between <sun> vs. <son> is 1 dB, while that of <mums> vs. <moms> is 2 dB; both are below the JND of 3 dB.

Intensity	**Sun**	**Son**	**Mums**	**Moms**
Speaker 1F	68	66	64	69
Speaker 2F	68	72	66	65
Speaker 3F	61	63	63	61
Speaker 4F	66	67	66	65
Speaker 5F	69	71	61	66
Speaker 6F	54	55	41	49
Speaker 7F	79	75	75	76
Mean	**67**	**66**	**64**	**62**
St. Dev.	6	7	8	10

Table 15: Intensity Measurements of [ʌ] and [ɔ] in CMNE in Experiment 2

5.6 Masking Analysis of Duration Based on Experiment 2

The JND for duration changes depending on the overall length of the speech signal under consideration. For signals whose duration is ≥ 200 ms, the JND is 20 ms, and for those with a duration greater than 300 ms, the JND is 30 ms.

Duration	**Sun**	**Son**	**Mums**	**Moms**
Speaker 1F	243	165	360	368
Speaker 2F	217	242	395	342
Speaker 3F	131	210	277	292
Speaker 4F	294	257	373	313
Speaker 5F	134	163	367	282
Speaker 6F	242	214	306	452
Speaker 7F	206	238	346	311
Mean	**212**	**209**	**337**	**346**
St. Dev.	37	59	58	41

Table 16: Duration Measurements of [ʌ] and [ɔ] in CMNE in Experiment 2

The temporal distance between [ʌ] and [ɔ] in <sun> vs. <son> is only 3 ms. That between [ʌ] and [ɔ] in <mums> vs. <moms> is also only 9 ms; both are considerably smaller than the required JNDs of 20 ms and 30 ms. This means that the two vowels mask each other in duration. This finding is very important because it suggests that the participants in Experiment 2 did not rely on duration to try to differentiate between [ʌ] and [ɔ]. Since, they failed to rely on temporal cues, confusion will be even greater because all other correlates mask each other in Experiment 2.

6.0 Discussions of Experiment 1

The analyses in the previous sections have focused exclusively on arithmetic means. When one takes such a bird's eye view, what statisticians call the "tyranny of averages" happens and skews the data for all the participants. To avoid this, a granular inspection is required so that we can focus on individual talkers in Experiment 1. We do so by considering the data displayed in Figures 1 and 2:

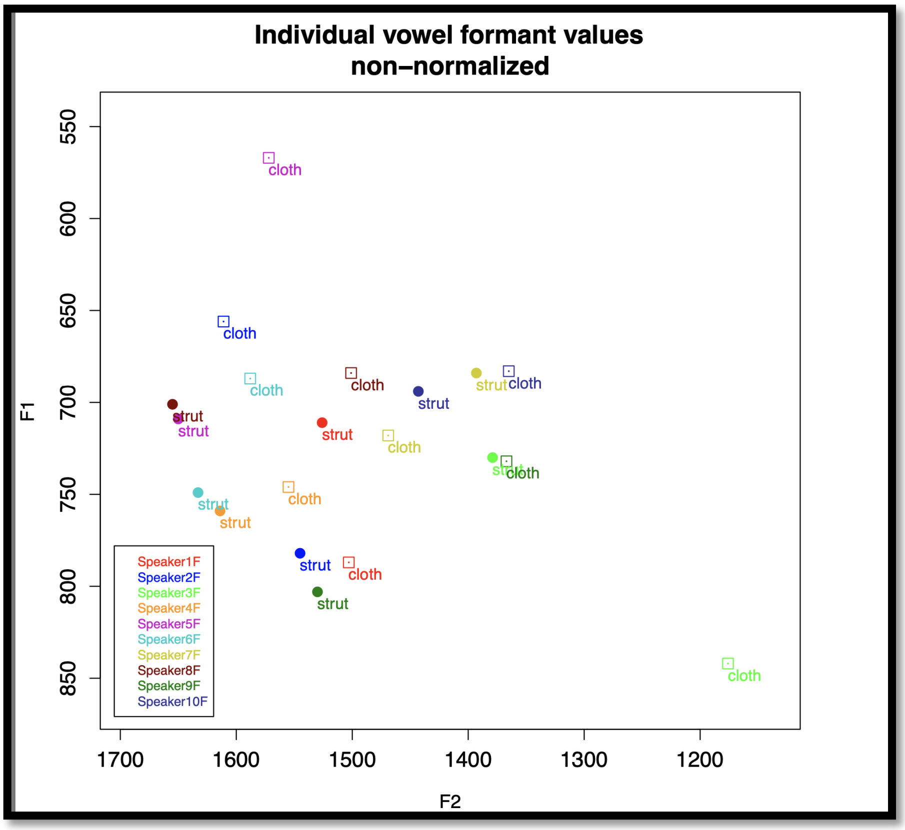

Figure 1: Masking by Individual Female Speakers in Experiment 1[6]

Among the female participants, masking occurs in the speech of only 4 individuals: Speaker 4F, 7F, 8F, and 10F. Three of them, Speaker 4F, 8F, and 10F, do not make any difference at all between [ʌ] and [ɔ] on the F1 frequency bandwidth because the acoustic distance between their vowels is less than 20 Hz (Thomas 2011:56). We also see that when Speaker 7F produce [ʌ] (684 Hz), it masks [ɔ] produced by Speaker 10F (683 Hz). Similarly, Speaker 3F produces [ʌ] (718 Hz) the same way Speaker 7F produces [ɔ] (718 Hz). There are several other patterns of interspeaker masking that are overlooked for the sake of brevity. We note in concluding this section that Speaker 3F produced [ʌ] and [ɔ] without any masking in F1 or F2. The fact that her F2 is 203 Hz is an indication that her accent is slightly different from the rest of the female participants.

[6] In the acoustic vowel space, "strut" refers to the vowel [ʌ] and "cloth" [ɔ]. These are the names by which sociolinguists refer to these two vowels. Also, this diagram and the others are produced using NORM.

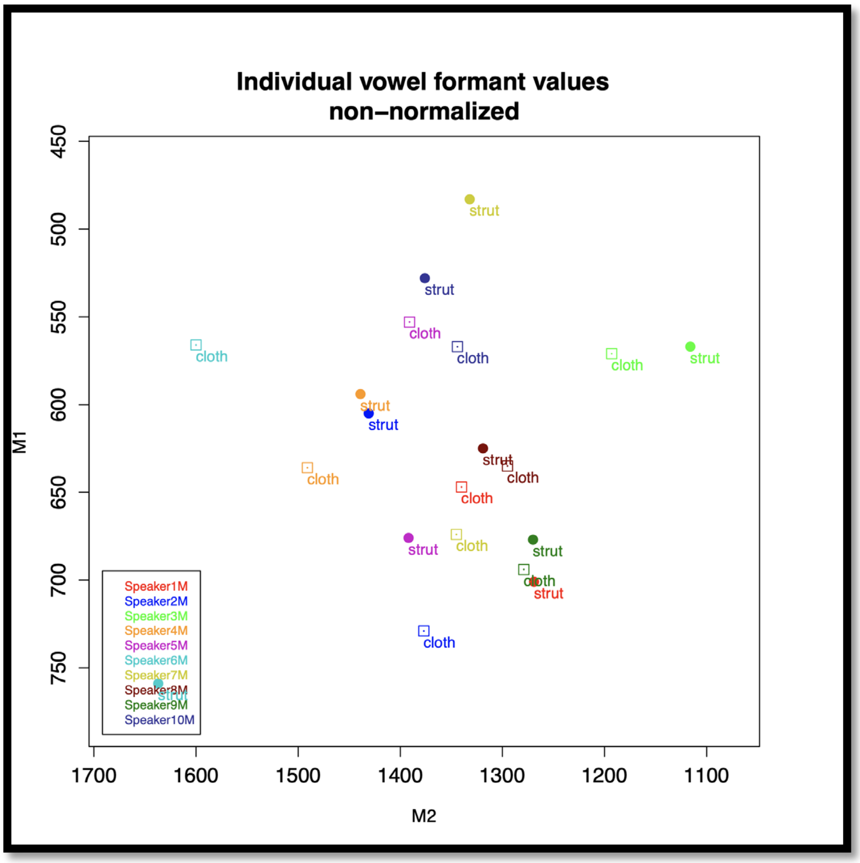

Figure 2: Masking by Individual Male Speakers in Experiment 1

Masking between [ʌ] and [ɔ] is more pronounced among male talkers. It is found in the speech of seven of the 10 participants, namely, Speakers 1M, 4M, 6M, 8M, 9M, and 10 M. Masking is absolute for 4 speakers: 3M, 8M, and 9M. To summarize, in Experiment 1, [ʌ] and [ɔ] mask each other completely for 30% of the participants, partially for 25% of the participants, for a total masking rate of 55%.

6.1 Discussions of Experiment 2

Experiment 2 is subdivided into two: the pronunciation of [ʌ] and [ɔ] in <sun> vs. <son> and in <mums> vs. <moms>. Figure 3 focuses on the former, and Figure 4 on the latter. The arithmetic means skewed the data more strongly in Experiment 2 than in Experiment 1. Even though the overall F1 mean is 49 Hz, masking occurs only in the pronunciation of Speakers 1F and 2F for <sun> vs. <son>. The remaining speakers, 3F, 4F, 5F, 6F, and 7F produced [ʌ] and [ɔ] distinctly, as shown in Figure 3:

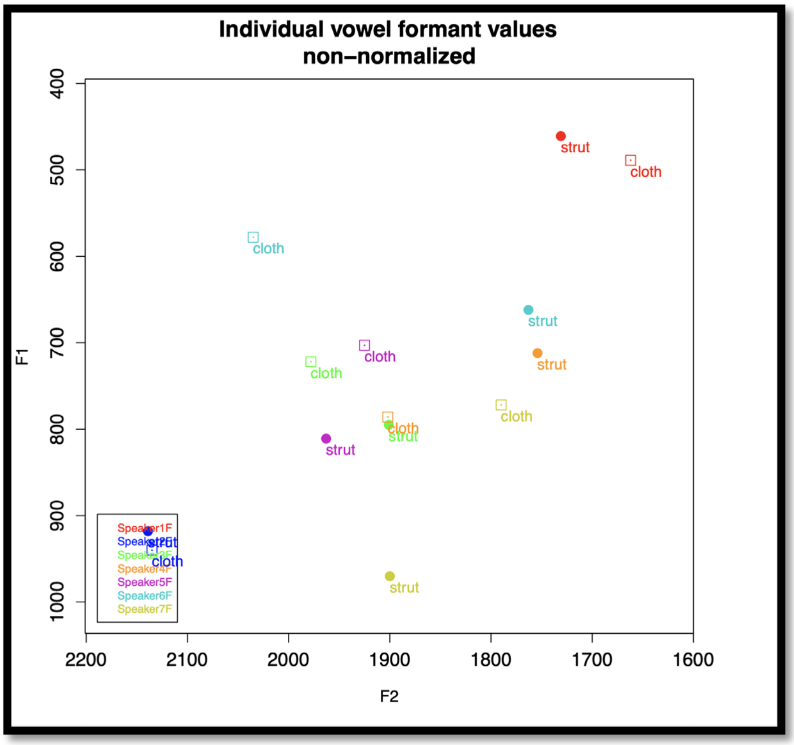

Figure 3: Masking Levels of <Sun> vs. <Son> Experiment 2

The pronunciation of [ʌ] and [ɔ] in <mums> vs. <moms> is displayed in Figure 3. Here only Speakers 1F and 6F produced them similarly. All the others distinguished between them clearly.

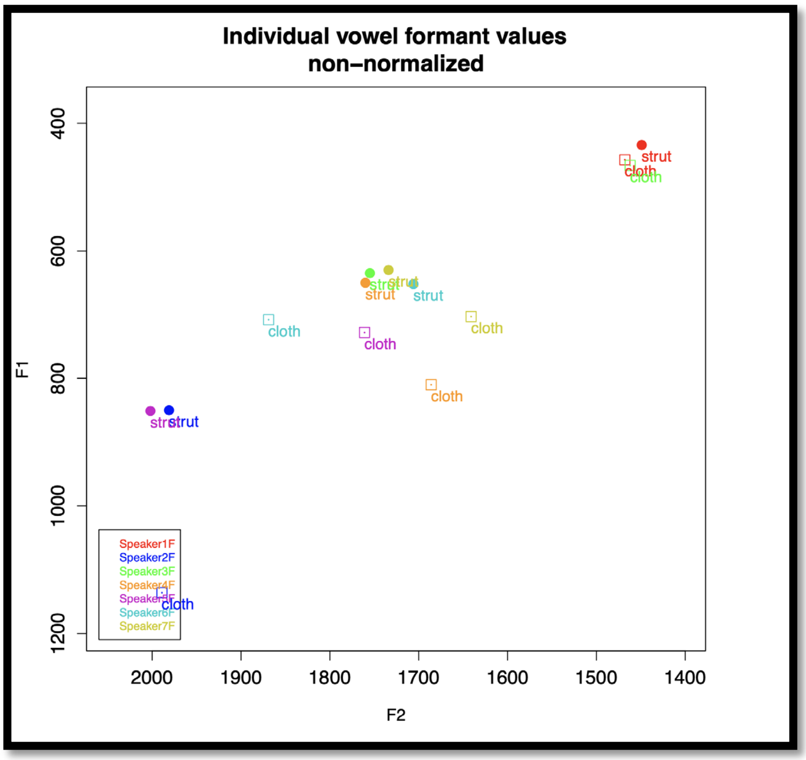

Figure 4: Masking Levels of <Mums> vs. <Moms> Experiment 2

Interspeaker masking occurs. When 3F produces [ɔ], its F1 of 466 Hz can be confused with [ʌ] produced by Speaker 1F (461 Hz). Only Speaker 1F displayed intraspeaker consistency. Speaker 2F exaggerated her pronunciation of <moms> so that it would sound different from <mums>. She opened her mouth wide and as a result she produced a very low [ɑ] whose F1 is 1136 Hz! Speaker 2F produced an [ɑ] sound for both <sun> and <son>, because her F1s are respectively 918 Hz and 940 Hz. Speaker 4F did the same, but in a less exaggerated fashion. Speaker 5F produced an [ɑ] (850 Hz) sound in <mums> instead of [ʌ]. Speaker 7F did the same with <sun> by producing [ɑ] instead of [ʌ].

Upon closer examination of the measurements in Experiment 2, we see that when CMNE monitor their pronunciation, in an effort to differentiate between [ʌ] and [ɔ], they produce [ɑ] instead. As a result, they substituted [ɔ] for [ɑ] twice (Speakers 2F and 4F) and [ʌ] for [ɑ] twice (Speakers 2F and 7F). It seems that the go-to vowel for some CMNE talkers who want to differentiate between [ʌ] and [ɔ] is [ɑ]. This is not entirely surprising because [ɑ] has overtaken [ɔ] in CMNE in many phonological environments except before liquids. It, therefore, makes sense for some to pronounce <son> and <moms> as [sɑn] and [mɑmz] because [n] and [m] are nasal stops. As for the substitution of [ʌ] by [ɑ], it is not entirely clear why Speakers 2F and 4F are produced <sun> as [sɑn], unless they resorted to this "faked" pronunciation to mark a clear distinction between <that sun is bright> and <that son is bright>. The participants in Experiment 1 who did not know what Author 2 was researching did not resort to such substitutions.

7.0 Summary

Twenty-seven talkers from Central Minnesota participated in two acoustic phonetic experiments designed to verify whether or not the pronunciation of [ʌ] and [ɔ] mask each other. The aggregated data in both Experiments 1 and 2 confirms that masking occurs consistently in F1, F2, F3, F4, and intensity, and inconsistently in duration. The only correlates in which [ʌ] and [ɔ] do mask each other are F0. This is not surprising because English is not a tone language. Masking in F3 and F4 is not as important as masking in F1 and F2 because higher formant frequencies correlate with speaker anatomical idiosyncrasies whereas masking in lower frequencies affect intelligibility. The fact that masking occurs in F2 shows that acoustically speaking, CMNE is a homogeneous dialect. Masking in F1 has a greater effect on speech intelligibility. When this correlate is considered, 14 out of the 27 participants (51.85%) produce [ʌ] and [ɔ] identically in <sun> vs. <son> and <mums> and <moms>. It is therefore not surprising that the girls on the Swim and Dive team were confused when the Booster Club president asked them to pick up the <mums>. Vowels are changing all over the country and in Central Minnesota. We consider ourselves fortunate to witness this change in progress.

ABOUT THE AUTHOR

Ettien Koffi, Ph.D. linguistics, teaches at Saint Cloud State University, MN. He is the author of five 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 perspective 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), Text-to-Speech (TTS), Speech-to-Text (STT), Intelligent Systems, and voice biometrics for speaker verification. He can be reached at enkoffi@stcloudstate.edu.

Connor Hanzsek-Brill is an undergraduate student at Saint Cloud State University, MN. He is responsible for the data collection and analysis used in Experiment 1. He often works with/for Dr. Koffi in his *Introduction to Linguistics* course as a tutor. He has an interest in East African and Eurasian languages, constructed languages, and phonetics. He hopes to specialize in phonetic research, linguistic field research, and community awareness programs that promote multilingualism. He can be reached at connorhanzsek@gmail.com.

References

Cao, Honglin and Volker Dellwo. 2019. The Role of the First Five Formants in Three Vowels of Mandarin for Forensic Voice Analysis. International Conference on Phonetic Sciences. https://icphs2019.org/icphs2019-fullpapers/pdf/full-paper_796.pdf.

Catford, J. C. 1987. Phonetics and the Teaching of Pronunciation. In J. Morley (Ed.), *Current Perspectives on Pronunciation: Practices Anchored in Theory*. Washington, DC: Teachers of English to Speakers of Other Languages, pp. 87-100.

Delattre, Pierre. 1951. The Physiological Interpretation of Sound Spectrograms *Proceedings of MLA*, Volume 66 (5):864-875.

Fromkin, V. R. Rodman, and N. Hyams. 2017. An Introduction to Language. Eleveth Edition. Boston, MA: Cengage Learning.

Heller, Eric J. 2013. *Why You Hear What You Hear: An Experiential Approach to Sound, Music, and Psychoacoustics*. Princeton, NJ: Princeton University Press.

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

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

Koffi, Ettien. 2021. Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility. Boca Raton: FL: CRC Press.

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

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

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

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. 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/.