Linguistic Portfolios

Volume 10 Article 3

2021

Vowel Intelligibility Analysis of Female Saudi Spoken English

Mahdi Duris
Iowa State University, mduris@iastate.edu

VOWEL INTELLIGIBILITY ANALYSIS OF FEMALE SAUDI SPOKEN ENGLISH

MAHDI DURIS

ABSTRACT

This article describes the acoustic characteristics of female Saudi-accented English vowels and uses acoustic phonetic measurements to assess the intelligibility of their vowels. Peterson & Barney's (1952) and Hillenbrand et al.'s (1995) methodology is slightly modified. Whereas their studies extracted various measurements, including F1 and F2 of vowels in General American English (GAE), the current study examines measurements extracted from 11 monophthong phonemic English vowels in running speech produced by 23 female Saudi EFL teachers. Intelligibility assessments of their vowels are based on Koffi's (2019) Acoustic Masking and Intelligibility (AMI) theory. He contends that intelligibility of vowels can be measured instrumentally by comparing the F1 of vowels because this formant carries 80% of the acoustic energy found in vowels. The AMI theory also combines Just Noticeable Differences (JND) thresholds and Relative Functional Load (RFL) calculations to gauge severity of masking and intelligibility. The findings in this article are based on 1,518 measured tokens. The study reveals that the fleece vowel [i], the kiss vowel [ɪ] and the face vowel [e] are the most problematic for Saudi speakers of English.

Keywords: Arabic-accented English, L2 Speech Intelligibility, Masking Analysis, Acoustic Phonetics of Saudi-accented English Vowels

1.0 Introduction

This article focuses on an acoustic phonetic analysis of the vowels of 23 female Saudi university teachers of English as a Foreign Language (EFL). By providing such an analysis, a clear picture is made available on the similarities and differences of Saudi-accented English vowels. Acoustic measurements allow us to target the features of an individual's speech intelligibility. In English, intelligibility can be increased by focusing on vowels. The academic consensus on vowels is best resumed by Prator & Robinett (1972:13): "if you wish to understand and be understood in English, you must be able to distinguish and make the distinction among the vowels sounds with great accuracy." First, Saudi spoken English (SSE) vowels are analyzed by mapping a clear picture of their acoustic characteristics. The study continues with a contrasting analysis of SSE vowels to General American English (GAE) vowels. This information highlights the differences in acoustic distances from both Saudis and Americans which are indicators for intelligibility. Finally, intelligibility assessments are performed to determine which vowels may cause poor intelligibility. These measurements allow us to propose a complete picture of the vowel intelligibility of Saudi spoken English.

2.0 Analytical Considerations

In the past, acoustic phonetic analyses of this type relied on Peterson and Barney's (1952) and Hillenbrand et al.'s (1995) studies of vowels in isolation, also known as citation form (CF). Unlike the two previous studies, this one will focus on vowels in running speech (RS). In so doing, we arrive at a more realistic, everyday classroom type of speech. As Saudi learners of English acquire vowels that are non-existent in Arabic (Al-Eisa, 2003), it is expected that some unfamiliar English vowels would also lead to poor intelligibility, but others may not (Khalil, 2014). Finally, to assess intelligibility, this study follows the rigorous *Acoustic Masking and Intelligibility (AMI)* theory set forth by Koffi (2019:73)[1] 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.

The AMI theory offers a robust acoustic phonetic measurement of intelligibility, which departs from the impressionistic assessment often used in L2 pronunciation studies. Fletcher (1953:153) describes masking as follows: "If while a sound A is being impressed upon the ear, another sound B is gradually increased in intensity until the sound A can no longer be heard, the sound A is said to be masked by the sound B. The sound A will be called the 'maskee' tone and the tone B the 'masker' tone." Masking in the frequency domain plays a very important role in intelligibility of speech segments, but especially vowels, which is the topic of this paper. In L2 pronunciation, as proposed by Koffi (2019), there are two types of masking: internal masking and external masking. The first deals with whether or not speakers' own vowels mask each other in pronunciation. The second seeks to determine if the speakers' vowels overlap with native speakers' vowels in auditory space.

2.1 Vowel Inventory

English vowels have important features that can be potentially difficult for L2 speakers to acquire. Fogerty & Humes (2012:1492) contend that "vowels are a primary element of the syllable (i.e., nucleus)." In order to highlight fundamental differences and similarities between Arabic and GAE, we will first look at their vowel inventories, starting with Arabic. Several proposals have been made with regard to Arabic vowel quadrants. However, Mansour Alghamdi, an acoustic phonetician from the Kingdom of Saudi Arabia (http://www.mghamdi.me/) who has been studying them for more than 20 years, has proposed the one in Table 1.

	Front	Central	Back
High	i: (Yaa)		(Wauw) u:
	i (Kasra)		
		(Damma) u	
Mid			
		(Fatha) a	
		(Alif) a:	
Low			

Table 1: Arabic Six Vowel Phoneme Model

Saudi speakers have only six vowels in their first language (L1). It is, therefore, a steep curve when they begin learning to speak English since General American English (GAE) has 11 phonemic vowels. English has five more vowels that Saudi Arabic does not have. As mentioned by Koffi (2019:90), the repository of cross-linguistic phonological inventory data, also known as PHOIBLE

[1] The classroom version is cited here and in the rest of the paper. However, a published version of the book is now available as Koffi, Ettien. 2021. *Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility*. Boca Raton, FL: CRC Press.

has inventoried 266 languages having systems of 3 to 9 vowels. Acquiring eleven vowels is a challenge for non-native speakers, specifically for Arabic natives. Table 2 gives a summary of all eleven phonemic GAE vowels.

	Font	Central	Back
High	/i/ <see>		/u/ <sue>
	/ɪ/ <sit>		/ʊ/ <soot>
Mid	/e/ <say>		/o/ <soak>
	/ɛ/ <set>		/ɔ/ <salt>
Low		/ʌ/ <such>	
	/æ/ <sat>		/ɑ/ <sod>

Table 2: GAE Eleven Vowel Model (Koffi, 2019:9)

Peterson and Barney (1952) and Hillenbrand et al. (1995) gave precise measurements of vowel formants (F1 and F2) which are essential for an accurate account of intelligibility. Ladefoged & Johnson (2015:221) describes these formants and the information they give us: “Spectrograms are usually fairly reliable indicators of relative vowel quality. The frequency of the first formant certainly shows the relative vowel height quite accurately. The second formant reflects the degree of backness quite well.” For this reason, these two formants will play an important role in assessing the intelligibility of Saudi spoken English.

2.2 Intelligibility Assessment

The methodology used to assess intelligibility for this study will replicate much of Koffi’s (2019) work. We will first note that F1 is the most important formant used in assessing intelligibility. Secondly, the relevant acoustic threshold, called Just Noticeable Difference (JND) will be used to make an inventory of problematic vowels. Lastly, we will gauge intelligibility by calculating distances between vowels and using that information in tandem with Relative Functional Load (RFL) calculations. The two help to determine the severity of unintelligibility.

2.2.1 The Importance of Vowel Height (F1)

For Fogerty and Humes (2012:1490), vowels are particularly important for intelligibility, as highlighted by this statement: “The data suggest that the acoustic information present during vowels is essential for speech intelligibility.” Ladefoged and Johnson (2015, p. 207) concur and single out F1 as most prominent because it alone contains 80% of the energy in a vowel. Therefore, focus will be given to F1 as a measure of intelligibility.

2.2.2 Just Noticeable Difference (JND) Threshold

The thresholds at which intelligibility measurement are salient have been summarized by Koffi (2019:92) and pertain to the acoustic distance needed between two phonemes. Koffi (2019:92) explains that Labov et al. (2013) “have used the acoustic threshold of 60 Hz as a robust acoustic criterion for distinguishing between perceptually similar vowels.” Any distance ≥ 60 Hz means that no masking occurs and “intelligibility is optimal.” If the distance between two vowels is ≤60 Hz, then "masking is likely." Furthermore, "complete masking occurs when the F1 distance between two vowels is ≤ 20 Hz" (Koffi, 2019:93). For this study, we pay attention only to instances of complete masking, that is, when the acoustic distance between two vowels is ≤ 20 Hz.

2.2.3 Relative Functional Load (RFL)

In order to give a complete assessment of intelligibility, Relative Functional Load (RFL) measurements will also be used. The RFL table presented by Koffi (2019:67) details 54 vowel phonemes pairs and their RFLs. For example, the vowels [i] vs. [ɪ] have an RFL of 95%. This means that if a speaker fails to distinguish between them in pronunciation, in 95% of the cases, this confusion will result in severe unintelligibility. The RFL of the vowels [ʊ] vs. [o] is only 12%. This means that masking between these two vowels affects intelligibility minimally, if at all. Simply stated, the higher the RFL, the greater the likelihood of unintelligibility. If the vowels of "beat" and "bit" mask each other, misunderstanding will be greater than if the vowels of "pull" and "pool" mask each other.

2.2.4 Internal and External Masking

Distinguishing between internal masking and external masking is also important in assessing intelligibility (Koffi, 2019:94). As noted earlier, internal masking pertains to assessing how speakers pronounce their vowels. If two vowels produced by the speaker are not distant from each other by at least ≥ 60 Hz, it means that the speaker does not distinguish the two vowels in his/her own pronunciation. This is called an internal masking. External masking is when the acoustic distance between a speaker's vowels and a hearer's vowels is ≤60 Hz. Table 3 below summarizes the main ingredients for assessing intelligibility:

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

Table 3: Intelligibility Assessment Matrix (Koffi, 2019:93)

3.0 Methodology

The participants for this study are 23 female Saudi adults whose age ranges from 19 to 45 years old. All of them reside in the capital city of the Kingdom of Saudi Arabia (KSA), Riyadh. All female participants are EFL teachers at the world's largest female-only university. For all participants, vowel measurements were extracted from a modified elicitation paragraph (Appendix 1). Vowels in running speech are preferred to the traditional approach of vowels in citation form because, in running speech, vowel production approximates speech found in real life. Vowels produced in citation form (isolated speech) presents two disadvantages for L2 participants. The lack of context for a word might prove difficult for L2 speakers. Specifically, from the list of isolated words used in past studies, we find: <hod>, <hawed>, <who'd>, <hud>, and <hoed>. Secondly, the phonemic knowledge needed to produce vowels in citation form accurately is sometimes a dauting task even for native L1, let alone L2 speakers of English. For instance, many native speakers of English stumble over <hawed> and misread it. By using running speech, target words appear in context, which increases naturalness of speech. Naturalness is achieved through the elicitation paragraph in Appendix 1 which contains only frequently used words.

3.1 Data Analysis

PRAAT is the software used for the analysis and measurements. Each vowel analyzed was extracted from three words found in the elicitation paragraph. The extraction was made manually to maximize accuracy. Table 4 below shows the list of words containing the relevant vowels used in this study. Wells' (1982) lexical set is used for the name of the vowels.

Vowel sound and name										
fleece	**kiss**	**face**	**dress**	**trap**	**lot**	**thought**	**goat**	**foot**	**goose**	**strut**
[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]
Text equivalent										
please peas meet	with thick is	maybe faked paper	yellow edge red	ask pad mat	Bob dog frog	for bought corner	old go zone	good book cookie	blue scoop zoo	rubber duck must

Table 4: Vowel Sound Names from the Elicitation Paragraph

In total, measurements of 11 vowels produced 3 times for 23 female participants were analyzed. The first step in the analysis consisted of splicing the extracted vowel audios into one single audio file corresponding to the vowel sound. Then, spectrographs were created from which F0, F1, F2, F3, F4, intensity and duration measurements were extracted. All in all, the data for the analysis consisted of 5,313 tokens (11 vowels x 3 repetitions x 23 participants x 7 correlates). Only F1 and F2 were used for this article for a total of 1,518 measured tokens.

4.0 Vowel Height Characteristics of Female SSE

Table 5 displays the full set of data obtained from the 23 participants.

Vowel Sound	**fleece**	**kiss**	**face**	**dress**	**trap**	**lot**	**thought**	**goat**	**foot**	**goose**	**strut**
F1 Correlate	**[i]**	**[ɪ]**	**[e]**[2]	**[ɛ]**	**[æ]**	**[ɑ]**	**[ɔ]**	**[o]**[3]	**[ʊ]**	**[u]**	**[ʌ]**
KSAF1	423	480	504	551	866	739	552	560	532	431	731
KSAF2	389	480	516	637	1045	851	654	598	493	484	869
KSAF3	510	557	529	631	879	814	724	622	546	553	802
KSAF4	508	550	561	705	923	828	720	616	558	446	818
KSAF5	490	567	567	654	888	738	656	620	566	551	779
KSAF6	442	523	464	615	808	725	657	543	531	466	662
KSAF7	416	540	577	653	935	851	779	660	582	479	862
KSAF8	407	492	415	557	757	739	581	577	499	434	677
KSAF9	424	533	573	689	927	729	587	617	533	523	817
KSAF10	425	516	486	621	881	750	546	593	467	459	754
KSAF11	414	555	548	610	924	737	669	651	532	470	686
KSAF12	459	523	558	650	876	817	727	625	553	492	732
KSAF13	434	546	482	673	982	841	671	575	516	496	844

[2] data taken from Hillenbrand et al. (1995)
[3] data taken from Hillenbrand et al. (1995)

KSAF14	462	591	551	692	953	846	707	596	529	458	832
KSAF15	461	563	484	593	896	761	604	615	539	454	755
KSAF16	415	529	668	669	1004	815	644	594	716	451	803
KSAF17	417	551	544	698	930	788	650	596	494	475	800
KSAF18	468	501	505	657	812	725	670	612	509	457	707
KSAF19	466	503	503	567	865	790	739	641	544	440	724
KSAF20	446	523	457	601	611	755	660	706	525	457	743
KSAF21	420	529	477	651	854	723	668	569	542	466	694
KSAF22	424	518	590	645	860	710	571	555	496	525	678
KSAF23	478	518	546	714	882	816	741	738	554	476	789
Mean	**443**	**530**	**526**	**641**	**885**	**778**	**660**	**612**	**537**	**476**	**763**
St. Deviation	32.7	28.1	54.4	46.1	87.5	48.2	63.8	45.9	47.4	34.1	62.7
P&B[4] (1952)	**310**	**430**	**536**	**610**	**860**	**850**	**590**	**555**	**470**	**370**	**760**

Table 5: Vowel Height Measurements for Female SSE

Vowel heights have three main levels. High, mid and low vowel positions. The results obtained from the Saudi teachers of English are as follows: Eight vowels out of 11 (72%) have a standard deviation below 60 Hz. The *trap* [æ] vowel (87.5 Hz), the *thought* [ɔ] vowel (63.8 Hz) and the *strut* [ʌ] vowel (62.7 Hz) have the highest standard deviation. This deviation exceeds the JND threshold of 60 Hz for distinguishing between two different phonemes. The standard deviation of only three vowels (27%) is below 60 Hz. This means that the participants produced these vowels fairly similarly.

5.0 Internal Masking and Intelligibility

Intelligibility is assessed relative to the masking threshold that is under 20 Hz for the pairs of vowels in Table 6. If the acoustic distance between any two vowels is < 20 Hz, then internal masking occurs. Severity of unintelligibility takes into account RFL percentages.

Vowel Pairs	F1 Distance	Internal Masking Levels		RFL		Intelligibility Rating
[i] vs. [ɪ]	87 Hz	No masking		95%		Good intelligibility
[ɪ] vs. [e]	**4 Hz**	**Complete masking**		**80%**		**Poor intelligibility**
[e] vs. [ɛ]	114 Hz	No masking		53%		Good intelligibility
[ɛ] vs. [æ]	245 Hz	No masking		53%		Good intelligibility
[u] vs. [ʊ]	61 Hz	No masking		7%		Good intelligibility
[ʊ] vs. [o]	75 Hz	No masking		12%		Good intelligibility
[o] vs. [ɔ]	48 Hz	Slight masking		88%		Average intelligibility
[ɔ] vs. [ɑ]	118 Hz	No masking		26%		Good intelligibility
[æ] vs. [ʌ]	122 Hz	No masking		68%		Good intelligibility
[ʌ] vs. [ɑ]	**14 Hz**	**Complete masking**		**65%**		**Average intelligibility**
[æ] vs. [ɑ]	107 Hz	No masking		76%		Good intelligibility

Table 6: Internal Masking and Intelligibility of Female SSE Vowels

[4] stands for Peterson & Barney

The first instance of complete masking has to do with the *kiss* vowel [ɪ] (530 Hz) and the *face* vowel [e] (526 Hz). The acoustic distance between them is only 4Hz. With an RFL at 80%, the intelligibility is deemed poor. As an example, when a Saudi female speaker says <fit> and <fate>, no difference would be audible. This pair of vowels is difficult for the 23 participants to produce accurately. This presents a clear intelligibility issue. The second example of complete masking is the *strut* vowel [ʌ] (763 Hz) and the *lot* vowel [ɑ] (778 Hz), which are separated only by 14 Hz. With an RFL of 65%, this could lead to poor intelligibility. Figure 1 displays the pairs of vowels in which internal masking occurs:

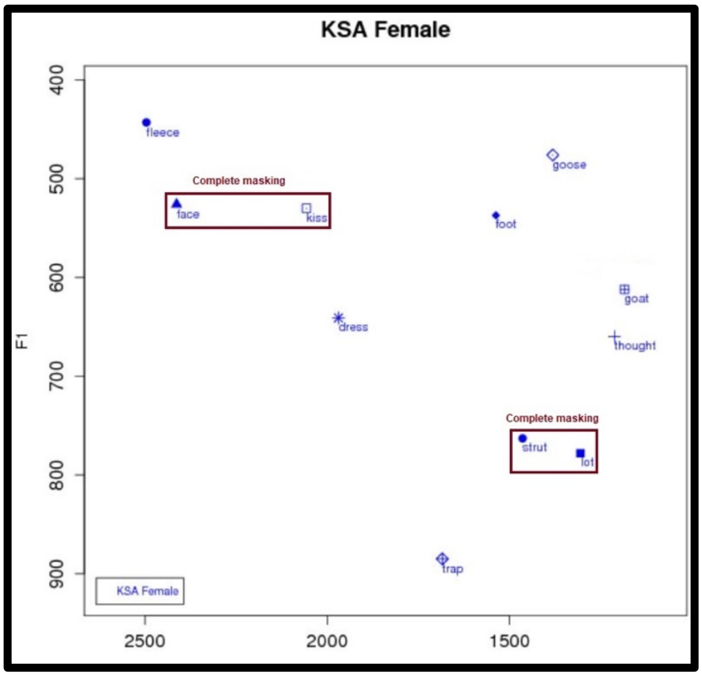

Figure 1: Internal Masking for Female SSE Vowels

6.0 Comparing Female SSE and Female GAE

In this section, Saudi speaker's vowels are compared and contrasted with those produced by GAE speakers. For F1, female SSE has 6 noticeable acoustic differences when compared to female GAE. The *fleece* [i], *kiss* [ɪ] and *goose* [u] vowels have the highest differences with noticeable difference above 60 Hz. Table 7 shows a complete comparison for vowel height.

Vowel Sound	**fleece**	**kiss**	**face**	**dress**	**trap**	**lot**	**thought**	**goat**	**foot**	**goose**	**strut**
Vowel	**[i]**	**[ɪ]**	**[e]**[5]	**[ɛ]**	**[æ]**	**[ɑ]**	**[ɔ]**	**[o]**[6]	**[ʊ]**	**[u]**	**[ʌ]**
					F1						
Female SSE	443	530	526	641	885	778	660	612	537	476	763
Female GAE	310	430	536	610	860	850	590	555	470	370	760
F1 difference	**133**	**100**	10	31	25	**72**	**70**	57	**67**	**106**	3

Table 7: F1 Data for Female SSE and GAE Vowels

From table 7, we see that Saudi speakers produce five vowels out of 11 (45%) similar to GAE speakers. For example, the *face* vowel [e] shows only an acoustic difference of 10 Hz. This means that the Saudi participants sound like GAE talkers when they say the *face* vowel. On the contrary, when Saudi participants say the *fleece* vowel, there is a strong difference with how GAE speakers say it. The vowel distance between these two groups of speakers for [i] is 133 Hz, way beyond the 60 Hz threshold. This happens for 6 vowels total out of eleven (55%). The next section will confirm if these differences in vowel height hinder intelligibility for Saudi spoken English.

7.0 External Masking and Intelligibility

External masking calculates the acoustic distance between vowels produced by Saudi speakers and those produced by GAE speakers. When the acoustic distance between two different speech segments is < 20 Hz, they externally mask each other completely, as shown in Table 8.

Vowel Pairs	**F1 Distance**	**External Masking Levels**		**RFL**		**Intelligibility Rating**
[i] vs. [ɪ]	**13 Hz**	**Complete masking**		**95%**		**Poor intelligibility**
[ɪ] vs. [e]	**6 Hz**	**Complete masking**		**80%**		**Poor intelligibility**
[e] vs. [ɛ]	84 Hz	No masking		53%		Good intelligibility
[ɛ] vs. [æ]	219 Hz	No masking		53%		Good intelligibility
[u] vs. [ʊ]	**6 Hz**	**Complete masking**		**7%**		**Good intelligibility**
[ʊ] vs. [o]	**18 Hz**	**Complete masking**		**12%**		**Good intelligibility**
[o] vs. [ɔ]	22 Hz	Moderate masking		88%		Poor intelligibility
[ɔ] vs. [ɑ]	190 Hz	No masking		26%		Good intelligibility
[æ] vs. [ʌ]	125 Hz	No masking		68%		Good intelligibility
[ʌ] vs. [ɑ]	87 Hz	No masking		65%		Good intelligibility
[æ] vs. [ɑ]	35 Hz	Moderate masking		76%		Average intelligibility

Table 8: External Masking and Intelligibility of Female SSE and Female GAE

[5] Hillenbrand et al. (1995)
[6] Hillenbrand et al. (1995)

GAE hearers are most likely to mistake Saudi-accented *fleece* vowel [i] (443 Hz) for *kiss* vowel [ɪ] (430 Hz) and vice versa because the acoustic distance between them is only 13 Hz. When female Saudi speakers produce the *fleece* sound [i] (443 Hz), it masks the GAE *kiss* sound [ɪ] (430 Hz) completely. With an RFL at 95%, this makes it completely unintelligible. For example, if a Saudi speaker says <seat>, it will be misperceived as <sit> by a GAE hearer. The *goose* vowel [u] (476 Hz) and *foot* vowel [ʊ] are also likely to be misperceived vowels. When Saudi speakers produce [u] (476 Hz), GAE hearers may perceive it as [ʊ] since the acoustic distance between them is 6 Hz. When Saudi speakers say <pool>, it may be heard by GAE speakers as <pull>, and vice versa. However, because the RFL is very low at 7%, this external masking is unlikely to cause intelligibility problems. The third instance of complete masking occurs with the *kiss* vowel [ɪ] (530 Hz) which masks completely the GAE *face* vowel [e] (536 Hz), with only 6 Hz of separation. Since their RFL is 80%, this masking causes serious intelligibility problems. When the participants say <fit>, it can be perceived as <fate> by GAE hearers. The fourth example of external complete masking is between the *foot* vowel [ʊ] (537 Hz) in SSE and the *goat* vowel [o] (555 Hz) in GAE. The acoustic distance between them is only 18 Hz. However, since their RFL of 12% is low, intelligibility issues are minimized. A poignant real-life example for external masking consequence is described by Koffi (2019:71), when he shares a personal story with a salesperson at an electronic store:

> I told the seller that I was looking for <i-pads> and wanted to take a look at what he had. The clerk was a native GAE hearer, and I am not. He disappeared for a moment and came back loaded with all types of <i-pods>. Why did he mistake my pronunciation of <i-pad> for <i-pod>? This happened because my [æ] and [ɔ] masked each other. The F1 of my [æ] is 820 Hz and that of my [ɔ] is 812 Hz. The salesperson mistook my [æ] for [ɔ] because the acoustic distance between my two vowels is only 8 Hz instead of the minimum 20 Hz required to avoid absolute masking.

Similar to this confusion, figure 2 below depicts such errors between the speech of Saudi speakers and GAE hearers.

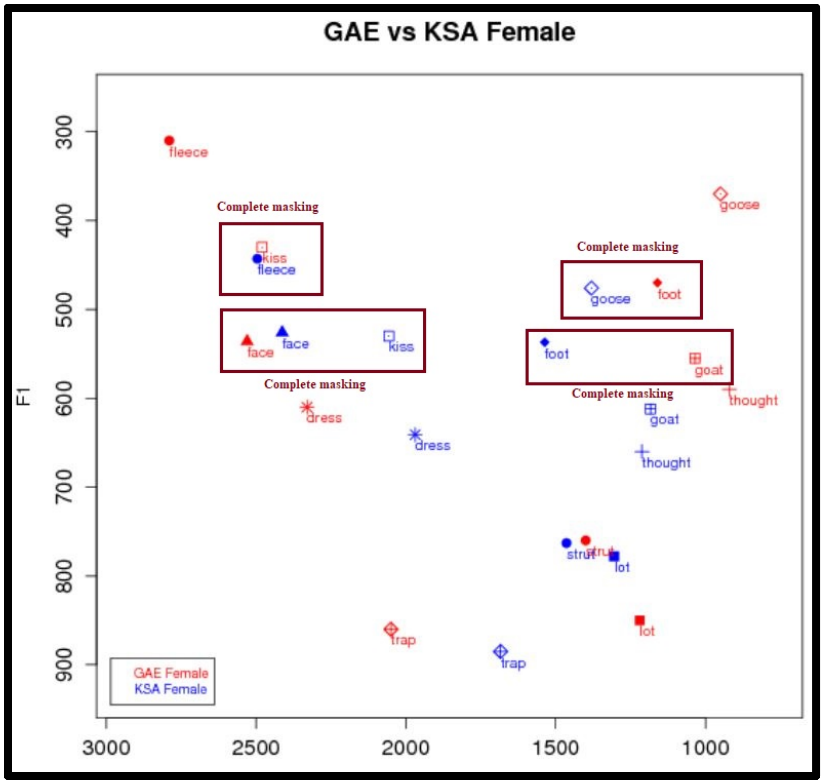

Figure 2: External Masking for Female SSE and GAE Vowels

Overall, GAE hearers are likely to misunderstand Saudi-accented English if the discourse context is not redundant enough. Saudi-accented [i] can be easily confused with [ɪ] by GAE hearers, and so can [ɪ] for [e], [u] for [ʊ], [ʊ] for [o].

8.0 Summary

The vowel space shows that the 23 female EFL teachers produce 8 out of 11 English vowels (81%) accurately. Of the three remaining vowels, the *kiss* vowel [ɪ] (530 Hz) and the *face* vowel [e] (526 Hz) mask each other internally. The *strut* vowel [ʌ] (763 Hz) and the *lot* vowel [ɑ] (778 Hz) also mask each other. Internal masking between these vowels interferes with intelligibility because their RFLs are high. External masking analysis reveals that Saudi-accented [ɪ], [u], [ʊ] can cause varying degrees of intelligibility problems because they overlap with GAE vowels in auditory space. Based on the findings discussed in this paper, pedagogical attention should be given to these vowels when teaching English to Saudi speakers. Greater attention should be focused on making sure that the *fleece* vowel [i] is produced distinctly from the *kiss* vowel [ɪ], and that it also is distinguished from the *face* vowel [e] since these three vowels have very high RFL, 95% and 80% respectively. The *goose* vowel [u] and the *foot* vowel [ʊ] on the one hand, and the *foot* vowel [ʊ] and the *goat* vowel [o] also need attention, but not as urgently as the vowels [ɪ, ʌ, ɑ].

ABOUT THE AUTHOR

Mahdi Duris is a Ph.D. Applied Linguistics and Technology student at Iowa State University. He holds a B.S. in Management from George Mason University (2006). After teaching financial literacy in Spanish for 6 years, he then obtained a CELTA certificate from Cambridge University (2012) to teach English as a Second Language to adults. He taught English in the Kingdom of Saudi Arabia for 5 years. His research focus includes pronunciation, acoustic phonetics and technology assisted L2 learning. He can be reached at mduris@iastate.edu.

Acknowledgments: I owe a debt of gratitude to Dr. Koffi for insightful comments on the pre-publication version of this paper.

References

Al-Eisa, A. (2003). *An ESL/EFL teacher's guide to English-Arabic interlingual phonology.* Thesis (M.A.). St. Cloud State University.

Fletcher, H. (1953). *Speech and Hearing in Communication.* New York: D. Van Nostrand Company, Inc.

Fogerty, D., & Humes, L. (2012). The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences. *The Journal of the Acoustical Society of America, 131*(2), 1490-1501.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic Characteristics of American English Vowels. *Journal of the American Acoustical Society, 97*(5), 3099-3111.

Khalil, S. (2014). *Comparative Study of the Acoustic Vowel Space of Egyptian English vowels and General American English vowels.* Thesis (M.A). St. Cloud State University.

Koffi, E. (2019). *Relevant acoustic phonetics of L2 English: Focus on Intelligibility.* St. Cloud, MN: Course Manuscript.

Ladefoged, P., & Johnson, K. (2015). *A course in phonetics (7th ed.).* Stamford, CT, USA: Cengage Learning.

Peterson, G. E., & Barney, H. L. (1952). Control Methods in a study of the vowels. *The Journal of the Acoustic Society in America, 29*(2).

Prator, C. H., & Robinett, B. W. (1972). *Manual of American English Pronunciation, 3rd ed.* New York: Holt, Rinehart and Winston.

Wells, J. C. (1982). Accents of English. Volume 1: An Introduction (pp. i–xx, 1–278), Volume 3: Beyond the British Isles (pp. i–xx, 467–674). Cambridge University Press. ISBN 0-52129719-2, 0-52128541-0.

Appendix 1[7]: Elicitation Paragraph

Note: The vowels in bold and capital letter are those that were measured acoustically in this study.

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 thick 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 also need a small plastic snake, the little y**E**llow b**OO**k, a r**U**bber 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 f**A**ked 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 sc**OO**p these things into three r**E**d bags and to **O**ld backpacks. We will g**O** meet her, Sue, Jake, and Jenny, Wednesday at the very last train station. The station is between the bus stop and the c**OO**kie store on Flag Street. We m**U**st m**EE**t there at 12 o'clock, for sure. The entrance **I**s at the **E**dge of the z**OO** in z**O**ne 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.

[7] This is an augmented version of the Speech Accent Archive text found at https://accent.gmu.edu/. The original text lacked the "foot" vowel [ʊ]. Furthermore, some segments had severely limited distributions. The expanded version remedies these insufficiencies.