Koffi, E. (2017) An acoustic phonetic account of the confustion between [1] and [1] in seven varities of L2 English: Focus on intelligibility and accentedness. In M. O'Brien & J. Levis (Eds). *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Calgary, AB, August 2016 (pp. 73-84). Ames, IA: Iowa State University.

AN ACOUSTIC PHONETIC ACCOUNT OF THE CONFUSION BETWEEN [1] AND [1] IN SEVEN VARIETIES OF L2 ENGLISH: FOCUS ON INTELLIGIBILITY AND ACCENTEDNESS

Ettien Koffi, St. Cloud State University

The lateral [1] and the rhotic [1] are classified as liquids because their places and manners of articulation overlap in many languages (Ladefoged & Maddieson 1996, p.185). As a result, when some L2 speakers of English produce them, these two segments are perceptually indistinguishable to some speakers of English. This is likely to cause unintelligibility for hearers of General American English (GAE). Other L2 speakers produce them clearly and distinctly. Intelligibility is not compromised, yet their speech is perceived as heavily accented. Both pronunciation patterns are investigated acoustically to determine thresholds at which unintelligibility or accentedness is perceived. The participants in this study are 10 Arabic, 10 Japanese, 10 Korean, 10 Mandarin, 11 Slavic, 6 Somali, and 10 Spanish L2 speakers. Their pronunciations of [1] and [1] are compared and contrasted with those produced by 10 native speakers of American English. All 77 participants read the same Speech Accent Archive text. The acoustic correlates studied in this paper are F3 and vibration rates. Catford's (1987) relative functional load (RFL) data is used to determine the rate of intelligibility, while Ladefoged (1996)'s trilling threshold is used to gauge the degree of accentedness of the [1]s produced by non-native speakers.

INTRODUCTION

Lindau (1985, pp. 157-168) wrote a paper entitled *The Story of /r/* in which she investigated the phonetic properties of /r/ in four European and seven non-European languages. I follow her methodology but focus instead on the story of /r/ in seven varieties of L2- English. This paper is divided into four sections. The first tabulates the confusion errors resulting from the substitutions of [1] by [1] and vice versa. The second provides acoustic measurements used to possibly account for the confusion. The third discusses the sociolinguistic significance of the confusion. The fourth suggests kinematic exercises that can help improve intelligibility on the one hand, and reduce strongly-accented [1]s on the other hand.

QUANTITATIVE ANALYSIS OF THE CONFUSION BETWEEN /I/ AND /r/

Seventy-seven participants were examined in this study. Ten are native speakers of General American English (GAE). The remaining 67 participants include 10 speakers of Arabic, 10 Japanese speakers, 10 Korean speakers, 10 Mandarin speakers, 11 speakers of South Slavic languages/dialects (5 Croatians and 6 Serbians), 6 Somali speakers, and 10 Spanish speakers. All of them read the same text from the George Mason University's Speech Accent Archive. The text has 16 words containing the grapheme <r>

them occur prevocalically, seven of which were clusters: <bring, from, fresh, brother, frog, three, red, train> and eight others postvocalically: <her, her, store, for, her, brother, for, her>. The text also has five prevocalic <l>s: <please, slabs, blue, plastic, Stella> and four postvocalic <l>s: <call, also, small, will>. Together the 67 non-native speakers attempted 1,072 [J]s and 603 [I]s, for a total of 1,675 liquids. The quantitative analysis focuses on the 871 prevocalic laterals [I] and rhotics [J]. The remaining 804 segments are not discussed because they occur in syllable codas. It is well known that in this position, liquids are often vocalized or even deleted (Lindau, 1985, p. 157 and Ash, 1982). The quantitative analysis of the narrow phonetic transcription shows that 833 liquids were produced accurately. In 33 instances [I] and [J] were confused with each other. In five other instances, they were misperceived as [n] or [w], as shown in Table 1.

Table 1

	Perceived Stimuli					
Spoken		[1]	[n]	[1]	[w]	Total
Stimuli	[1]	309	3	23	0	335
	[I]	9	0	525	2	536

Confusion of Matrix of /l/ and /r/

The total erroneous pronunciations amount to 4.36%, but the actual confusion errors attributable to prevocalic [l] and [J] represent 3.78% of the data. A closer examination of the data in Table 2 shows that Japanese, Mandarin, and Korean talkers made most of the errors:

Table 2

Confusion	Rate	by	Language	Group
-----------	------	----	----------	-------

	Confusion Rate	Percentage
Arabic	0 out of 130	0
Japanese	23 out of 130	17.69%
Korean	6 out of 130	4.61%
Mandarin	7 out of 130	5.38% ¹
Slavic	1 out of 143	.69
Somali	0 out of 78	0
Spanish	0 out of 130	0
GAE	0 out of 130	0

Speakers from these three language groups produced 37 of the 38 errors (97.36%). The Japanese talkers alone made 23 errors (60.52%), followed by Mandarin talkers who made seven errors (18.42%), and by Korean speakers who made six errors (15.78%). However, Mandarin speakers confused [1] and [1] only twice. The five remaining errors have to do with three instances when [1] was transcribed as [n], and twice when [1] was perceived as [w]. Since [1] was confused with [1] only twice, we will consider it an incidental error. We will focus the rest of the paper on the errors made by Japanese and Korean speakers.

FEATURE DISTRIBUTION OF /l/ AND /r/

Why did GMU transcriptionists transcribe [1] and [1] as they did in Japanese and Koreanaccented English? Why didn't they make similar perception errors with the other accented Englishes? The answers may lie in the information presented in Table 3 regarding the kinds of liquids found in the native languages of the speakers. The data in the second and third columns are taken from Ladefoged and Maddieson (1996, pp.185, 216). That in the last column is based on information from the *Handbook of the International Phonetic Association* (1999) and Maddieson (1984).

¹ The higher confusion rate of Mandarin-accented English requires some explanations. There were only two instances where [l] and [r] were confused with each other. However, in three instances, [l] was mispronounced as [n], and in two others, [r] was produced as [w]. The actual [l] and [r] confusion rate is 1.53%. Since this paper focuses narrowly on [l] and [r], we will not address the confusion between [l] and [n], nor will we discuss the one between [r] and [w].

Table 3

N0	Articulatory Descriptions	IPA Symbols	Languages
1.	Voiced alveolar lateral	1	Mandarin, Slavic, Somali
2.	Voiced laminal dental lateral	1	Spanish
3.	Voiced apical alveolar lateral	1	Arabic
4.	Voiced apical post-alveolar	1	
5.	Voiced laminal post-alveolar	1	
6.	Voiced lateral alveolar flap	1	Korean
7.	Voiced alveolar or alveolar trill	r	Arabic, Mandarin, Slavic, Somali, Spanish
8.	Voiced dental or alveolar tap or flap	ſ	GAE, Korean
9.	Voiced dental or alveolar approximant	T	GAE, Spanish
10.	Voiced post-alveolar flap	t	Japanese
11.	Voiced post-alveolar approximant	ન	
12.	Voiced uvular trill	R	
13.	Voiced uvular approximant	R	
14.	Voiced dental or alveolar lateral flap	I	

Articulatory Descriptions of /l/ and /r/

Two rhotics are listed for Spanish because it has both an approximant flapped [I] and a trilled [r] (Lindau 1985, p.161). Korean also has two liquids. The lateral /[/ is the basic phoneme, but it has [r] as its allophone in intervocalic positions (Lee 1999, pp.120-123). The status of /r/ in Mandarin is in limbo. Some take it to be an allophone of /l/, but the existence of minimal pairs between [l] and [I] (Smith 2010, pp. 88-89) suggests that [I] may be a phoneme in its own right. Okada (1999, p. 118) indicates that the Japanese postalveolar rhotic [t] occurs mainly between vowels. Arabic, Slavic (Croatian and Serbian), and Somali have only trilled [r]s. Except for Japanese, all the languages have a lateral segment.² The distributional restrictions on liquids in Japanese and Korean may explain why the pronunciations of [l] and [I] were transcribed with the other sound more than the other groups of L2 English speakers. However, since our interest in this paper is not phonology, we turn our attention to acoustic phonetics for explanations.

ACOUSTIC PHONETIC ACCOUNTS

Before providing the relevant acoustic measurements on which the explanations of the analyses will be based, we must discuss briefly the methodology and equipment used. The recordings were downloaded from the Speech Accent Archive website. A commercial software package called "Wav Pad" was used to convert the videos into audio files and save them as .wav files. The words <Stella> and <red> were selected to study the acoustic characteristics of [l] and [J] produced by all 77 participants. Each file was annotated in Praat (Boersman and Weenink 2012, Version 5.3.16) using Ryan's (2005) Grid.maker.praat. Five tiers were created to study F1, F2, F3, Intensity, and Duration. The spectrographs of <red> produced by MN 143M illustrates the procedure that was followed for all the speakers:

² There are disagreements on the IPA symbols used for Japanese and Korean liquids. Some use [t] or [I] for Japanese. Others use [t] or [r] for Korean rhotics.



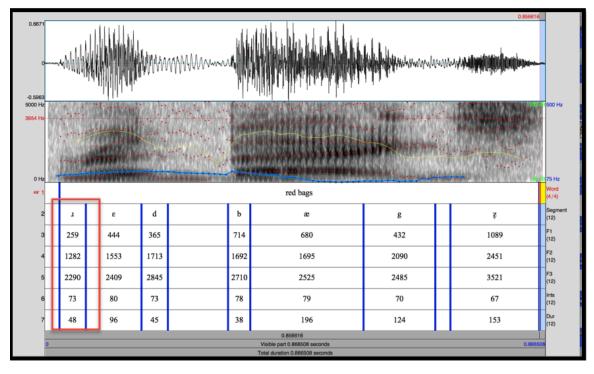


Figure 1. Spectrograph of <red> in <three red bags> by MN 143M

Yoon's (2008) script, Stress-analysis.praat, was used to collect all measurements. A rectangle has been drawn around [1] to draw attention to it since all the other measurements in the spectrogram are irrelevant to this analysis.

F3 Measurements of L2-accented [l]s and [J]s

The data was compiled for each of the 77 participants. The measurements were then averaged per language group, as shown in Table 4.

Table 4.

F3 of [l]	and [1]
-----------	---------

	F3 of [l]	F3 of [1]
Arabic	2,619	2,397
Japanese	2,779	2,674
Korean	2,754	2,697
Mandarin	2,708	2,451
Slavic	2,738	2,517
Somali	2,943	2,483
Spanish	2,869	2,495
GAE	2,827	2,225

The acoustic correlate that phoneticians deem the most robust for discriminating between [1] and [1] is F3 (Lindau 1985, p. 158). It goes without saying that it is the correlate used in this study. Espy-Wilson's (1992:739, 747) study of [1] and [1] in GAE is considered the most authoritative to date. She measured their occurrences in three phonological environments: prevocalic, intervocalic, and postvocalic. The mean F3 measurements for [1] are, respectively, 2,553 Hz, 2640 Hz, and 2630 Hz. The mean in all three positions is 2,607 Hz. The measurements for [1] are as follows: 1,779 Hz prevocalically, 1,720 Hz intervocalically, and 1,830 Hz postvocalically. Its mean F3 in all positions is 1,776 Hz. These measurements are in line with other studies. Generally, the F3 range for [1] is from 2,600 to 3,000 Hz, while that for [1] goes from 1,200 Hz to 2,000 Hz.³ Ladefoged and Maddieson (1996, p.244) explain that "a lowered third formant is a well-justified specification for the various articulations of the American I."

Cross-linguistic studies of [l] and [I] show similar measurements. Except for apical/dental trills, rarely does one find a rhotic whose F3 is 2,600 Hz. This suggests that any liquid whose F3 reaches 2,600 Hz is very likely to be categorized as a lateral. On the basis of available data on the F3 of liquids, Koffi (2016) has proposed the Liquid Intelligibility Criterion (LIC):

Liquid Intelligibility Criterion (LIC)

[I] masks [l] if its F3 is \geq 2,600 Hz, unless it is trilled.

³ The mean of all GAE talkers in Koffi (2016) is 2,225 Hz. This is likely due to the fact that all the [r]s in the data occur before the front vowel [ε].

The criterion implies that [I] with F3 less than 2,600 Hz may also be confused with [J]s. However, F3 alone is not a necessary and sufficient condition to predict that [J] and [I] will mask each other. The escape clause "unless it is trilled" makes it clear that trilling is the difference. Ladefoged and Maddieson (1996, p.203) have reported that there is no known lateral trill in any human language. Consequently, any liquid that is trilled is automatically perceived as a rhotic. Six Japanese talkers: Japanese 2F, 4M, 9M, 11F, 12M, 13M, and seven Koreans: 1M, 2F, 4F, 8F, 9M, 10M, 12M, 13F, produced [J]s with F3 values at or beyond 2,600 Hz.⁴ Did they trill them? If they did, no confusion would arise. If they did not, confusion is predicted to occur.

Vibration Rates of L2-Accented [l]s and [J]s

Before answering the question from the previous section, let's see how trilling is calculated. According to Lindau (1985, p.166), the determination that a rhotic is trilled or flapped depends on its total duration. Ladefoged (1996, pp. 114-116) provides the following formula for calculating trilling:

Duration in milliseconds

Vibration Rate in Hz = _____

Segment duration Ref

A rhotic is perceived as trilled if its vibration rate is at or exceeds 22 Hz (Ladefoged and Maddieson 1996, p. 218, 226). This is an important threshold that can be used in tandem with F3 measurements to explain why GMU transcriptionists perceived some Japanese and Korean-accented [l]s as [I]s, and vice versa. The data in Table 5 provides some answers:

⁴ The explanations for these measurements found in Koffi (2016, pp. 258-264).

Table 5

	Duration in Ms		Vibration in Hz		Difference	
Words	Stella	red	Stella	red	Red - Stella	
Segments	[1]	[1]	[1]	[L]	[]-[L]	
Arabic Mean	73	40.9	13.59 Hz	24.44 Hz	10.85	
Japanese Mean	73.35	70.6	13.63 Hz	14.16 Hz	.53	
Korean Mean	74.8	70.1	13.36 Hz	14.26 Hz	.9	
Mandarin Mean	66	59.4	15.15 Hz	16.83 Hz	1.68	
Slavic Mean	65	46.9	15.38 Hz	21.32 Hz	5.94	
Somali Mean	76.6	33.1	13.05 Hz	30.21 Hz	17.16	
Spanish Mean	67	46.4	14.92 Hz	21.55 Hz	6.63	
GAE Mean	62	52.5	16.12 Hz	19.04 Hz	2.92	

Duration and Vibration Rates of [l] and [.]

The vibration rates indicate that, overall, Japanese and Korean talkers flapped their [1]s. Eight of the 10 Japanese participants and nine of the 10 Korean talkers did so. Only Japanese 8M, 10F, and Korean 11M trilled their [1]s. In other words, 17 out of 20 (85%) participants produced [1]s that are aurally indistinguishable from theirs [1]s. This explains why GMU transcriptionists transcribed most of their [1]s as [1]s.

MARKEDNESS IN THE PRONUNCIATION OF RHOTICS

Cross-linguistically, the phonetic realizations of rhotics have been found to have both regional and social dialectalization significance. Labov's (1972, pp. 43-69) "Fourth Floor" study in three New York City department stores has helped draw attention to this phenomenon. In L2 contexts, rhotics can elicit three types of responses from hearers, two of which are sociolinguistically marked, while one is unmarked (Wardhaugh and Fuller 2015, pp. 101-102). Let's consider them briefly.

Perceptually Marked Pronunciations Caused by Confusion

Failure to distinguish clearly between laterals and rhotics has intelligibility consequences. According to Catford (1987, p. 88), the relative functional load (RFL) between [I] and [I] in prevocalic positions is 83%. This high percentage can lead to serious unintelligibility. Furthermore, the existence of minimal pairs or near minimal pairs creates lexical competitions that can interfere with intelligibility. If a Japanese or Korean speaker substitutes [I] for [I] or vice versa in *
bring> vs. <bling>, <fresh> vs. <flesh>, <blue>*

vs. <brew>, <frog> vs. <flog>, and <red> vs. <led>, an interlocutor may have a hard time deciphering the intended utterance. As a matter of fact, Japanese 2F, 11F, 12M, and Korean 1M pronounced <red> as <led>, while Japanese 5F and Korean 8F pronounced <frog> as <flog>.⁵ It is true that the discourse context can mitigate the seriousness of the confusion, but as we all know, context is not a panacea against all misunderstandings.

Perceptually Marked Pronunciations Caused by Trilling

Trilling [r] is a double-edged sword. On the one hand, it enhances intelligibility, but on the other hand, it marks the speech as heavily accented. It enhances intelligibility because, as noted earlier, no trilled lateral has been found in any human language. Consequently, as soon as a trill is heard, English transcribers automatically perceived it as a rhotic. With regard to GAE, a trill [r] indexes the talker as having a foreign accent because in this dialect of English, the vibration rate of all rhotics is below 22 Hz. The trilling of [r]s in Arabic and Somali-accented Englishes caused GMU transcriptionists to perceive accurately all of the 128 rhotics that they produced. Yet, overall, their speech is heavily accented in part because of the strong trilled [r] that they produce. The Arabic-accented [r] is 2.44 Hz above the trilling threshold of 22 Hz. It should, however, be noted that three talkers, Arabic 1F, Arabic 35M, and Arabic 51M flapped their <r>s, while the seven remaining talkers forcefully trilled theirs. All six Somali talkers without exception trilled their [r]s vigorously, that is, 8.21 Hz above the threshold. For the speakers of these two languages, intelligibility is traded for a "heavy" or "thick" foreign accent.

Perceptually Unmarked Pronunciations

On average, the rhotics produced by Mandarin, Slavic, and Spanish talkers fall below the 22 Hz threshold. Sociolinguistically, speech items that do not attract accentedness attention are considered unmarked (Wardhaugh & Fuller, 2015, pp. 101-102). The vibration rates for these three groups of speakers may be surprising, since the rhotics in their respective native languages are classified as trills. My findings are in agreement with Lindau (1985, p. 161) who states that "Even in languages where a possible realization is a trill, not all speakers use a trill, and the speakers that do, have tap and approximant allophones as well as the trill. In the languages used in this study that were described as having an apical trill, about half of the speakers produced trills, but not for every token." She also observed that, in her study, most of the Spanish participants trilled their [r]s. My findings confirm hers. Of the 10 participants, Spanish 6M, 9M, 11F, and 13F flap their [r]s, while six others trilled theirs. The same is true for the 11 Slavic participants. Croatian 6F, Serbian 2M, Serbian 6M, and Serbian12F flapped their [r]s, while the seven others trilled theirs. The only group where the majority of the participants flapped their [r]s is Mandarin. Only two, Mandarin 6F and 12M, out of 10 trilled their [r]s.

⁵ Other segmental errors occurred, such as the pronunciation of [f] as $[\phi]$ or [p], but we are not concerned with such errors in this paper.

SUMMARY

The analyses presented in this paper show that unintelligibility and accentedness for [I] and [I] can be accounted for acoustically. The findings complement the impressionistic phonetic transcriptions made by the GMU linguists who transcribed the speech samples produced by the 67 non-native speakers who produced the lateral and rhotic segments under consideration. Reliance on well-established thresholds for the F3 of [I] and [I] has allowed us to provide an acoustically sound explanation for why GAE hearers have a hard time differentiating between the [I] and [I] produced by some Japanese and Korean L2 speakers of English. The vibration rate threshold for [I] also explains why the [r]s that Arabic and Somali talkers produce are indexed as heavily accented even though their pronunciation does not interfere with intelligibility. Lastly, even though <r> is trilled in many native languages, some speakers tend not to transfer the full force of their native [r]s into English. As a result, their L2-accented [I]s are unmarked.

ABOUT THE AUTHOR

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

REFERENCES

- Ash, S. (1982). *The vocalization of /L/ in Philadelphia*. Ph.D. Dissertation, University of Pennsylvania. Ann Arbor, MI: University Microfilms International.
- Boersman, P., & Weenink, W. (2012). *Praat: Doing phonetics by computer*. Available online at: <u>http://www.praat.org/</u>.
- Bradlow, Ann R., David Pisoni, Reiko Akahane-Yamada, and Yoh'ichi Tohkura. (1999). Training Japanese listeners to identify English /r/ and /l/: Perception and production. *Perception & Psychophysics* 61 (5): 977-985.
- Catford, J. C. (1987). Phonetics and the teaching of pronunciation. In J. Morley (Ed.), *Current Perspectives on Pronunciation: Practices Anchored in Theory* (pp. 87-100). Washington, D.C.: Teachers of English to Speakers of Other Languages.
- Espy-Wilson, C. (1997). Acoustic measures for linguistic features distinguishing the semivowels in American English. *Journal of the Acoustical Society of America* 92 (2), 736-757.

- International Phonetic Association. (1999). *Handbook of the International Phonetic Association*. New York: Cambridge University Press.
- Koffi, E. (2016). *Relevant acoustic phonetics of L2 English: Focus on intelligibility. Manuscript.* St. Cloud, MN.
- Ladefoged, P. (1996). *Elements of Acoustic Phonetics*, 2nd edition. Chicago: The University of Chicago Press.
- Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Malden, MA: Blackwell Publishers Inc.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Lee, H. B. (1999). Korean. In *Handbook of the International Phonetic Association* (pp. 120-123). New York: Cambridge University Press.
- Lindau, M. (1985). The story of /r/. In V. Fromkin (Ed.), *Phonetic Linguistics: Essays in Honor of Peter Ladefoged* (pp. 157-168). New York: Academic Press.
- Maddieson, I. (1984). *Patterns of sounds*. Cambridge Studies in Speech Science and Communication. New York: Cambridge University Press.
- NCH Software. (2013). *WavePad Sound Editor*, Version 5.17. NCH Software Inc.: Greenwood Village, CO, USA.
- Ryan, K. (2005). Grid-maker.praat. Available online at: http://www.linguistics.cla.du/faciliti/facilities/acoustic/praat.tml.
- Saito, K. & Lyster, R (2012). Effects of form-focused instruction and corrective feedback on l2 pronunciation development by Japanese learners of English. *Language Learning* 62 (2):595-633.
- Smith, J. G. (2010). Acoustic properties of English /l/ and /r/ produced by Mandarin Chinese speakers. MA Thesis: University of Toronto, Toronto, Canada.
- Wardhaugh, R. & Fuller, J. 2015. *An Introduction to Sociolinguistics*, 7th edition. Malden, MA: Wiley Blackwell.
- Weinberger, S. (2014). *Speech Accent Archive*. George Mason University. Retrieved from <u>http://accent.gmu.edu</u>.
- Yoon, T-J. 2008. Stress-analysis.praat. Available online at: http://web.uvic.ca/~tyoon/resource/vq.praat.

Appendix: List of Participants

Arabic Speakers: Arabic 1F, Arabic 30F, Arabic 35M, Arabic 36M, Arabic 40M, Arabic 44F, Arabic 46M, Arabic 47M, Arabic 50M, Arabic 51M

Japanese Speakers: Japanese 2F, Japanese 3F, Japanese 4M, Japanese 5F, Japanese 8M, Japanese 9M, Japanese 10 F, Japanese 11F, Japanese 12M, Japanese 13M

Korean Speakers: Korean 1M, Korean 2F, Korean 3F, Korean 4F, Korean 8F, Korean 9M, Korean 10M, Korean 11M, Korean 12M, Korean 13F

Mandarin Speakers: Mandarin 1F, Mandarin 2F, Mandarin 3M, Mandarin 4F, Mandarin 5F, Mandarin 6F, Mandarin 8M, Mandarin 9M, Mandarin 12M, Mandarin 19M

Slavic Speakers: Croatian 1F, Croatian 2F, Croatian 4M, Croatian 5F, Croatian 6F, Serbian 1F, Serbian 2M, Serbian 6M, Serbian 11M, Serbian 12F, Serbian 14F

Somali Speakers: Somali 1F, Somali 2M, Somali 3F, Somali 4F, Somali 5M, Somali 6M

Spanish Speakers: Spanish 1M, Spanish 2M, Spanish 4F, Spanish 6M, Spanish 9M, Spanish 11F, Spanish 13F, Spanish 14F, Spanish 16M, Spanish 20M