2024

ACCENTEDNESS AND INTELLIGIBILITY: THE TYPICAL PHONETIC STORY OF AN LX COUPLE IN THE US

MAHDI DURIS

ACCENTEDNESS AND INTELLIGIBILITY: THE TYPICAL PHONETIC STORY OF AN LX COUPLE IN THE US

MAHDI DURIS

ABSTRACT

For the past 20 years, intelligibility research has paved the way in making pronunciation teaching and learning about individual competency rather than sounding like a native speaker. The body of research has empirically shown that language speakers of English (Lx) can be intelligible while sounding heavily accented. However, studies have yet to examine how first-generation immigrant couples from different cultural backgrounds interact with each other using an Lx. A phonetic analysis of an oral transcription is conducted to describe the segmental features of Dominican English (DomE) and Pakistani English (PakE), two Lx English spoken by Tabinda and Tariq, as they recount their love story. Compared to General American English (GAE), results show that both speakers are intelligible and make expected substitution based on their first languages. The Dominican speaker substituted vowel [ɪ] for [i] and tapped rhotics (/ɾ/ for /ɹ/). For PakE, the most prominent feature is changing the stops' place of articulation using dental (/t̪/ /d̪/) and retroflex stops (/ʈ / /ɖ /). Additionally, some phonetic features are found to be specific to each speaker. Although some segmental features are not present in either couple's L1, the "redundancy" effect described by Koffi (2021:99) when quoting Fry helps facilitate intelligible speech between them. This study contributes to the growing body of research describing the speech features that make Lx English speakers intelligible.

Keywords: Accented-English Speech, Accentedness, Intelligibility, Lx Phonetics, Non-native English speakers, Sociophonetics.

1.0 Introduction

This article aims to present a summary of findings from an oral transcription and analysis of the spoken and phonetic features of two non-native speakers (NNS) of English. Tabinda and Tariq Sheikh shared the story that started their twenty-three years of marriage with NPR's *StoryCorps* podcast[1]. From this two-and-a-half-minute spoken interview, a segmental analysis of their respective Englishes has been conducted, explicitly focusing on their different types of L2 English. As Tabinda's native language (L1) is Dominican Spanish and Tariq's L1 is Urdu, they communicate using their L2-influenced English. This analysis highlights some of the most noticeable segmental features of Dominican English (DomE) and Pakistani English (PakE) in how they differ from General American English (GAE). Those features will be identified using PRAAT (Boersma & Weenink, 2007), and some broad conclusions will be made about Tabinda and Tariq's Englishes.

1.1 From L2 to Lx Terminology

Overall, this analysis seeks to shed light on a broader sociolinguistic reality for many first-generation immigrants who come to the United States for a different future. Amongst many challenges of leaving a home country, the social acceptance of the host country can be difficult to attain when one's primary communication resource, pronunciation, can be used by the autochthone

---

[1] The podcast is available for listening here: https://storycorps.org/stories/tariq-sheikh-and-tabinda-sheikh-161021/

to stigmatize because of one's accentedness. In the field of Applied Linguistics, successful work has been done by researchers (Munro & Derwing, 1995; Levis, 2005) to shift the pronunciation teaching paradigm from one focused on nativeness to one focused on intelligibility. Over 25 years of research have consistently shown that heavily accented speech can be highly intelligible to native listeners (Munro & Derwing, 2022). However, fewer research efforts have looked into how speakers from a different L1 communicate intelligibly using a common lingua franca that is different from their respective native languages. Furthermore, for research convenience purposes, most findings related to the phonetic features of Lx speakers gravitate around international college students, with fewer studies related to first-generation immigrants. Lastly, responding to the 2018 call from Dewaele, the common terminology L1/L2 will be replaced by L1/Lx to better align with the intent of this article. As previously mentioned, the field of applied linguistics has worked hard to move the needle from a native accent expectation in pronunciation to an intelligibility construct. Cohen (2002) has also proposed moving away from labels of nativeness attached to research, such as native speakers (NS) and NNS. Taking a Positive Psychology approach, Dewaele and Saito (2022:228) define Lx as any additional language acquired by an individual after age three. More relevant to this article, Dewaele and Saito (2022) offer the following observation for using an L1/Lx dichotomy:

> Also, the new L1/LX dichotomy is free of any ideological connotations and allows comparisons between L1 users and LX users that do not imply a deficit view. Moreover, the new dichotomy is holistic since multilinguals are L1 users as well as LX users. In other words, any judgment about the L1 and the LX is not about the whole person but only about some languages in their repertoire.

Lx has been adopted over LX in the literature to avoid confusion with the common short-hand use of LX for "linguistics." See Hayes-Harb et al., (2021); Beaulieu et al., (2022a); and Beaulieu et al., (2022b).

2.0 Lx Englishes
Tabinda and Tariq's story is very typical of many couples that come to live in the United States. Coined as a "melting pot," American society is well known worldwide for its diverse population. This diversity is also apparent in the mixed culture marriages that make up this melting pot. Like them, the author and his wife use specific types of Englishes as a typical "household" lingua franca. She is from Pakistan, and Urdu segmentals sometimes get in the way as they communicate. She is always open to replying to his mimics by highlighting how the author's French L1 gets in the way of his speech. But like them, what are these specific segmentals that most listeners recognize as accented speech but ignore immediately because intelligibility is high?

To analyze Tabinda's segmentals, two major categories were in focus: vowels and rhotics. Vowels in Dominican English, abbreviated as DomE could be a good indicator of specific features as Catalan Spanish (the variety of Spanish spoken in Latin America) uses a smaller inventory than American English (8 vowels vs. 12 vowels) as shown in Figure 1.
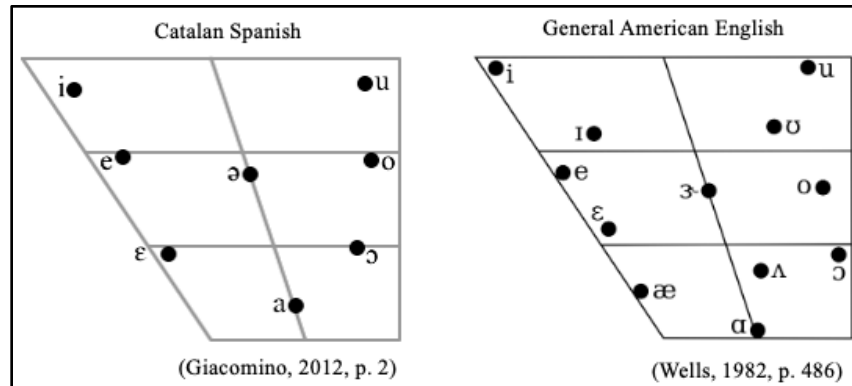
Figure 1: Vowel Space for Catalan Spanish and GAE

As Tabinda has been speaking DomE for more than 30 years now, we can expect that her vowel production is fossilized, and only a few vowels would differ from the local English variety. Rhotics between Spanish and English are also very distinguishable. Spanish uses taps and trills, and they have a contrastive intervocalic role in determining the meaning of a word (Willis & Bradley, 2008:87) as in *pero* (but) and *perro* (dog). Equally problematic, the approximant liquid /ɹ / in American English is perceived by many non-linguists as the most recognizable feature of American dialect.

For Tariq's segmentals, a focus was placed on several different features based on how Urdu influences the production of Pakistani English, known hereafter as PakE. The first one is the /w/ vs. /v/ contrast, followed by the retroflexion of /t/ and /d/, the dentalization of /t/ and /d/, the clear vs. dark /l/, and finally, some phonological considerations (epenthesis, rhoticity, and unaspirated stops). Mahboob and Ahmar (2008:1012) thoroughly explainall these features, which will guide the analysis of PakE.

2.1 Analysis Consideration
While analyzing the speech of Tabinda for specific features in DomE, some research in L2 English can inform us about what is already known from vowel and rhotic usage. Koffi (2021:98) confirms that some vowels can be problematic for L1 Spanish speakers. Specifically, differentiating between [ɪ] and [i], [ʊ] and [u], and [æ] vs. [a]. Willis and Bradley (2008) studied the taps and trills of Dominican Spanish and provided some important considerations. When differentiating between taps /ɾ/ and trills /r/, onset cluster and word-initial make the difference. A trill-only would be in word-initial as in *roca* (rock), *rojo* (red), and an "only tap" would occur after a consonant, as in *broma* (joke) and *gramo* (gram). In their study, they also provide mean measurements for differentiating between taps and trills, which were collected from Dominican female speakers (Willis and Bradley, 2008:95). All in all, when considering the speech of Tabinda, we will look at the features in Table 1:

	Vowels to consider		Tap	Trill
1.	[ɪ] and [i]	Length	> 21 ms	>85 ms
2.	[ʊ] and [u]			
3.	[æ] vs [a]	IPA Symbol	ɾ	r

Table 1: Vowels and Rhotic in DomE

For vowels in Tabinda's speech, we will not limit our analysis to the three mentioned; however, they will be a good guide when searching for vowel segment differences. For rhotic production, a segment will be identified as a tap and not an English /ɹ/ when the length is above 21 ms. Similarly, a trill will be distinguished from a tap when that segment is longer than 85 ms.

For Tariq's speech, the criteria in Table 2 are used to determine if a segmental is one specific to PakE or simply different from the local variety, GAE:

GAE	IPA		PakE	Symbol
Alveolar stops	[t], [d]	vs.	Retroflex stops	[ṭ], [ḍ]
Dental fricatives	[θ], [ð]	vs.	Dental stops	[t̪], [d̪]
Phonemic distinction between /v/ and /w/		vs.	**No phonemic** distinction in Urdu between /v and /w/. Depends on speaker choice.	
Dark vs. Clear [l]	/ɫ/, /l/	vs.	Only uses clear /l/	
Aspirated stops	/kʰ/, /pʰ/		Unaspirated stops	/k/, /p/

Table 2: Phoneme comparison between GAE and PakE
(Mahboob and Ahmar, 2008:1011-1013)

With these characteristics for DomE and PakE, we better understand which segmentals to focus on to highlight differences with GAE. The following section will describe how these will be further identified when using PRAAT.

2.2 Token Selection
Tabinda and Tariq's conversation was first transcribed into a word document.[2] Each utterance was given its line, and the speaker was identified. For each listening, a color note would be used for vowel differences in red and non-vowel differences in purple. When the speaker omitted a phoneme, they would be marked with an "x." Once I was clear about the utterances, PRAAT was used for data analysis. Seven tiers were created in a *TextGrid,* each serving a specific data collection function displayed in Figure 2. The first tier, "orthography," separated the dialogue between Tabinda and Tariq into their respective turn-taking utterances. In total, the recording generated a dialogue of 48 distinctive segments.

---

[2] Transcript available at https://storycorps.org/stories/tariq-sheikh-and-tabinda-sheikh-161021/
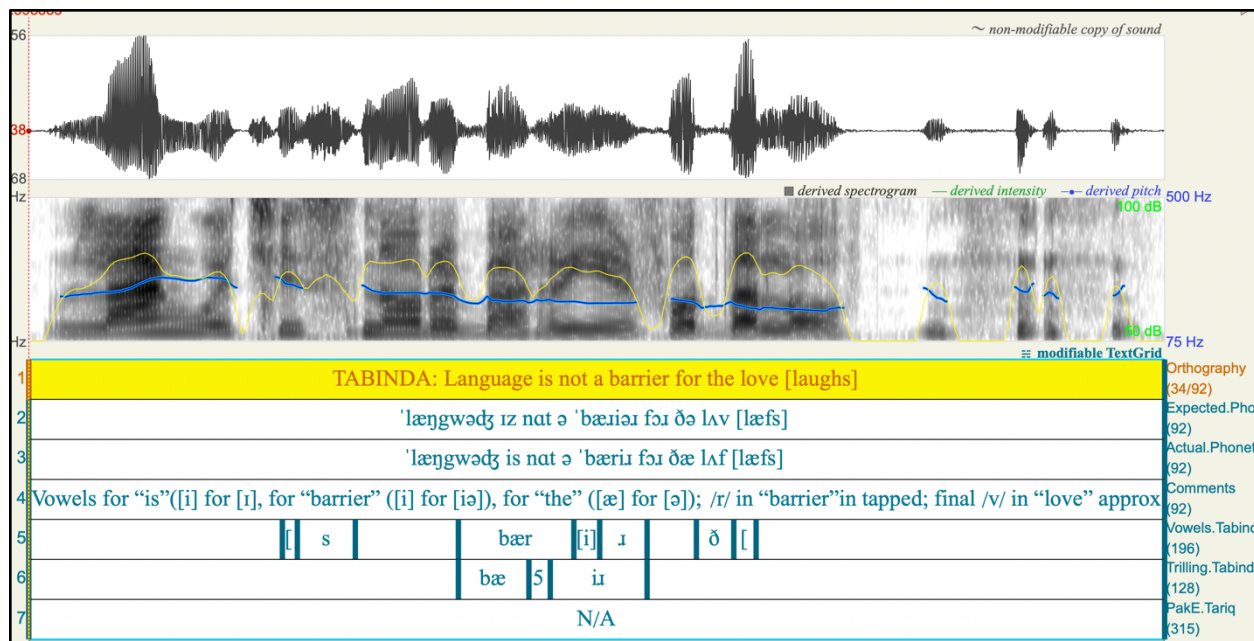
Figure 2: TextGrid for Segment 18

The next two tiers recorded the IPA of the "expected phonetic" transcription of the utterance, shown in Tier 2, while Tier 3 was that of the "actual phonetic" transcription of the utterance using IPA symbols. In total, 380 words were analyzed for Tier 3, a crucial step in identifying substitutions. Tier 4 was used to record "comments" about the segment in relation to segmental features. The last three tiers were used for segmental measurements. Tiers 5 and 6 measured the vowels and trills of Tabinda. Tier 7 measured the segmental features particular to Tariq. All vowels were measured using the spectrographic view by placing the cursor at the beginning of the darkening of striation from at least the three first formants. The whole vowel energy was selected, ending where the formants ceased to show apparent dark striations. Figure 3 depicts such procedure for Tabinda's production of "**the** love" where the schwa /ə/ was substituted for /æ/.
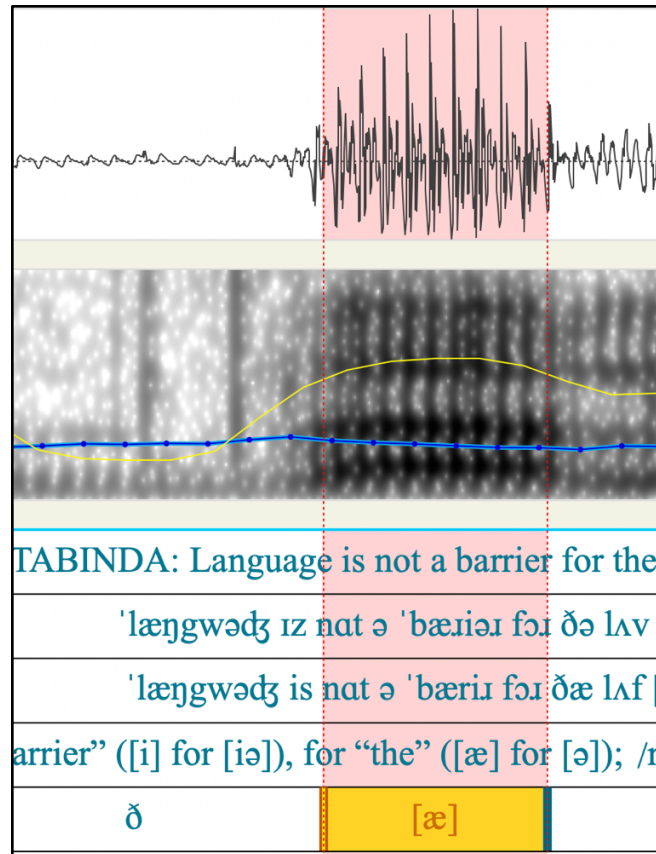
Figure 3: Measurement boundaries of /æ/ in Segment 18

For measuring rhotic features, the lack of striation in the spectrogram view would be the first indicator, placing a point of reference when a greyish striation started leading to the vowel in darker colors. For tap and trills, the repetitions in taping were selected until the next segment started while retaining the time duration measurement of that segment. For segment 18 below, the measurement in focus is for the word *barrier* which was produced using a trill of 51 ms after a heterosyllabic consonant.
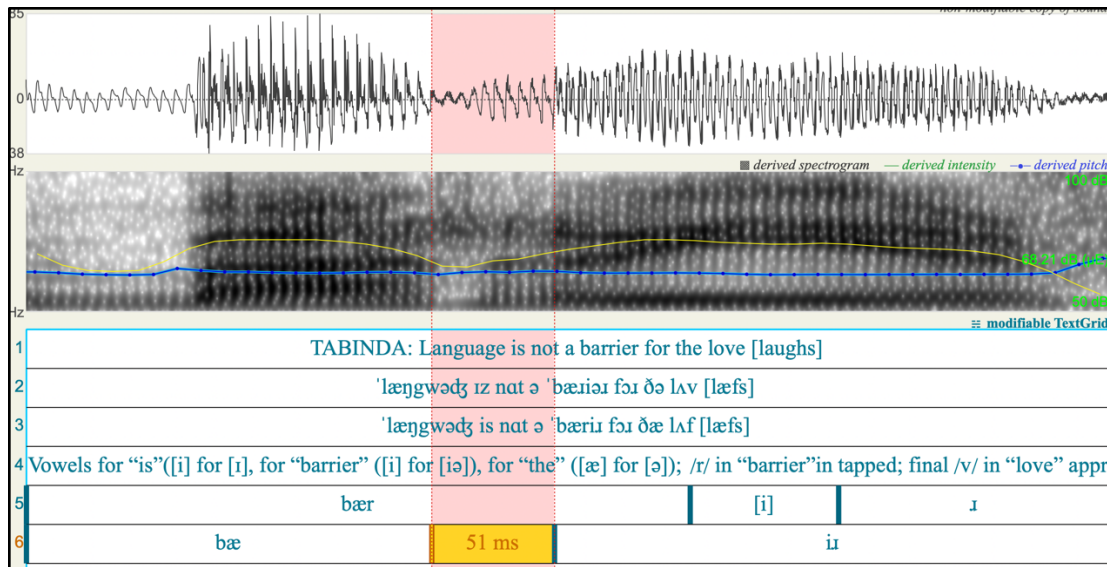
Figure 4: Measurement boundaries for trill /r/ in Segment 18

For other phonemes, the best method was trial and error in selecting the correct portions. For this analysis, Table 3 below shows the number of tokens analyzed for DomE.

	Vowel Tokens	Rhotic Tokens	Misc Tokens
Tabinda	24	16	21

Table 3: Summary of DomE Areas of Focus for Analysis

The miscelaneous category for Tabinda's production (i.e., 21 tokens) includes more than vowels and rhotics such as consonant substitutions. Table 4 details the tokens under analysis for PakE production.

/ð/ /θ/ vs. /t̪/ /d̪/	/t/ /d/ vs. /ʈ/ /ɖ/	/w/ to /v/ choice	/ɫ/, /l/	/pʰ/ /kʰ/ vs /p/ /t/
11	22	23	4	4

Table 4: Summary of PakE Areas of Focus for Analysis

3.0 Results
The results presented here are the most prominent features of each speech for DomE and PakE that differ from GAE. These speech features will also be considered for intelligibility since research has empirically shown that accented speech does not necessarily determine intelligibility (Levis, 2020). Furthermore, according to Koffi (2021, 47-50), L2 speakers' intelligibility depends on relative functional load (RFL) considerations rather than accentedness. First, the speech of Tabinda will be reviewed, and specific conclusions will be made according to that data. This will be followed by an analysis of Tariq's speech and a summary.

3.1 Analysis and Results of DomE by Tabinda
Tabinda uttered a total of 145 words, with 24 vowels that differed from GAE. For the most recurring vowel, she changed the "kit" vowel [ɪ] the most, using the the "fleece" vowel [i] instead. The second most occurring change happens when *the* is substituted for "da." In GAE speech, most instances for "the" are produced using a interdental fricative /ð/ followed by the lax-vowel schwa

/ə/; however, throughout this interview, Tabinda consistently substitutes this word by producing a voiced alveolar stop /d/ with a Spanish sounding vowel /a/ (i.e., da). As it relates to intelligibility, the prominent substitution of the KIT vowel /ɪ/ for the FLEECE vowel /i/ may bring some difficulties to some listeners to fully understand Tabinda. This vowel substitution ranks high on the RFL scale at 95%. For example, when Tabinda said, "language is /ɪz/ not a barrier," most GAE listeners would hear it as *language *ease* /iz/ not a barrier*.

For her production of rhotics, two tokens are not being counted because she used her L1 in context to say a Spanish word (*Gordito*). Four out of 16 of her rhotics were tapped /ɾ/ while only one was trilled with a length of 180 ms (*rude*). Most of the others are standard GAE /ɹ/ phonemes. Finally, looking at the other features in her utterances, some interesting production arises. The first relates to deleting /t/ when in the final position. She did so for *just, it,* and *thought.* The second is between the voiced and voiceless fricatives (/v/ and /f/). She replaced the voiced /v/ with the voiceless /f/ and said /lʌf/ (*love*) three times and /muf/ (*move*) two times. This substitution in fricatives is well known as coda devoicing, where the articulartory demands of producing a voiced sound (/v/) are challenging and replaced by a voiceless counterpart /f/. With that said, coda devoicing is also a feature of some GAE speakers as detailed in Koffi & Simmonds (2018) for Central Minnesotan English as an example.

3.2 Analysis and Results of PakE by Tariq
Tariq, the husband, uttered a total of 235 words, with 61 of those containing some specific features related to PakE. The first feature relates to how some speakers of PakE produce the GAE dental fricatives as dental stops mentioned by Mahboob et al., (2008:1011). In 11 occurrences of fricatives, Tariq made only 3 substitutions using dental stops for [d̪æt] "that, [t̪ɹi] " "three," and [vɪd̪] "with." This indicates that some of the features of GAE fricatives are still challenging for Tariq.

The second noteworthy feature from Tariq's speech are the substitutions of GAE alveolar stops for retroflex stops. This analysis uncovered ~~about~~ 20 retroflex substitutions, in either initial or final position and for either /ʈ/ /ɖ/. A retroflex stop differs from its alveolar counterpart in the positioning of the tongue being behind the alveolar ridge, generally defined as a more hollow sound. Figure 5 illustrates the differences in production from a alveolar stop /t/ and a retroflex stop /ʈ/ from Tariq's speech.
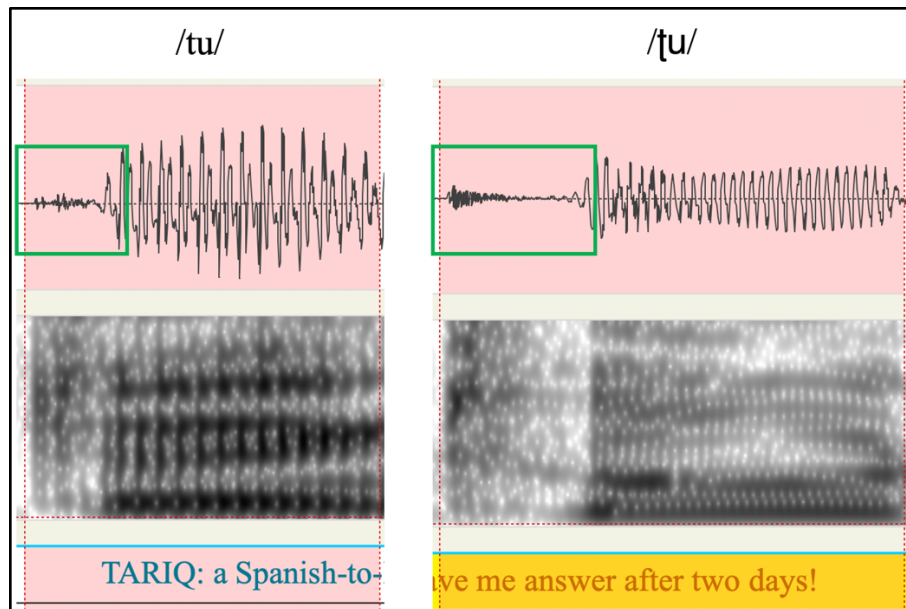
Figure 5: Comparision of an alveolar stop /t/ and a retroflex stop /ʈ/

The substitution of alveolar stops for retroflex stops is a distinguishable feature of Tariq's speech. Out of 39 words containing alveolar stops in initial or final positions, 22 of those words (56%) were substituted for retroflex stops as shown in Table 5. Additionally, some words contained substitutions in both word-initial and -final such as *today, tired*, and *don't*.

Substition type	Word initial position	Word final position
/t/ → /ʈ/	**t**wo, **t**ired, **t**ime, **t**elling, **t**ell, **t**oday,	seven**t**y, pu**t**, slep**t**, wha**t**, tha**t**, don'**t**, pocke**t**, bough**t**,
/d/ → /ɖ/	**d**on't, **d**reams, **d**ays, **d**rop,	tire**d**, ha**d**, coul**d**, to**d**ay,
Total	10	12

Table 5: Summary of alveolar substitutions for retroflex in Tariq's speech

The third PakE feature that is explored is between /v/ and /w/, which can be quite challenging, and Tariq does show this in his speech. Tariq produces three forms: the /v/ sound, the /w/ sound, and a mixed one that starts with a /v/ but morphs into a /w/ (noted /v/+/w/ in this dataset). All three sound types are visible with the spectrograms in Figure 6 with *where* (/**v**ɛr/), *what* (/**v**wɑt/), and *was* (/**w**ɑz/). Typical spectrogram features of the labiodental voiced fricate (/v/) are highlighted with a purple box, while features of the voiced labial-velar approximant (/w/) are shown in orange.
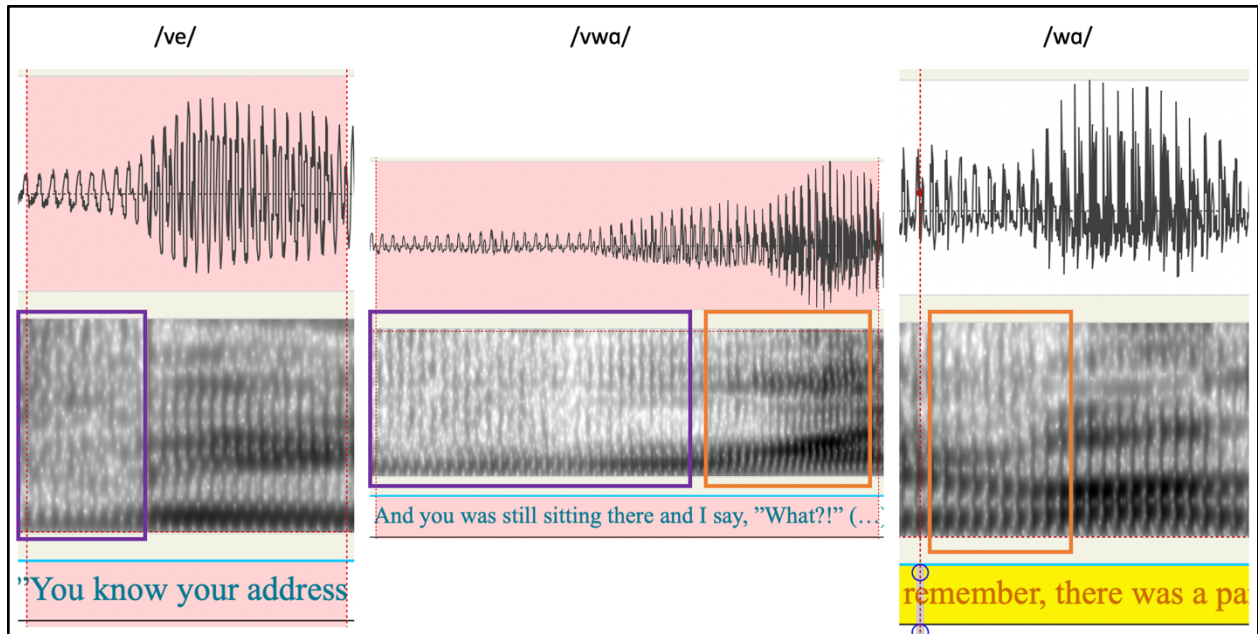
Figure 6: Spectrographic view of Tariq's *where, what*, and *was* production

Out of the 25 annotated, the main substitution type was replacing /w/ for /v/. All other occurences are presented below.

Substitution type	Token (# of occurences)	
/w/ → /v/	was (1)	YMCA (3)
	will (1)	why (1)
	with (2)	working (1)
	where (1)	
/w/ → /v+w/	was (2)	what (1)
	went (2)	
No substitutions	wanna (1)	when (1)
	was (6)	write (1)
	we (1)	

Table 6: Substitutions of /w/ by /v/

The fourth feature distinguishes between a dark /ɫ/ and a light /l/. Mahboob and Ahmar (2008:1011) state that Urdu does not have this distinction, where only the alveolar /l/ is used by PakE speakers. Although the conversation between Tariq and Tarinda only produced a limited number of laterals, ~~and~~ Tariq was able to produce ~~one~~ two dark /ɫ/ (*small* and *girl*) but mispronounced two as light ones (*still, will*) in line with Mahboob and Ahmar's findings. To better illustrate this feature, Figure 7 presents two utterances from Tariq's speech using the prevocalic /l/, here *like* and *live*.
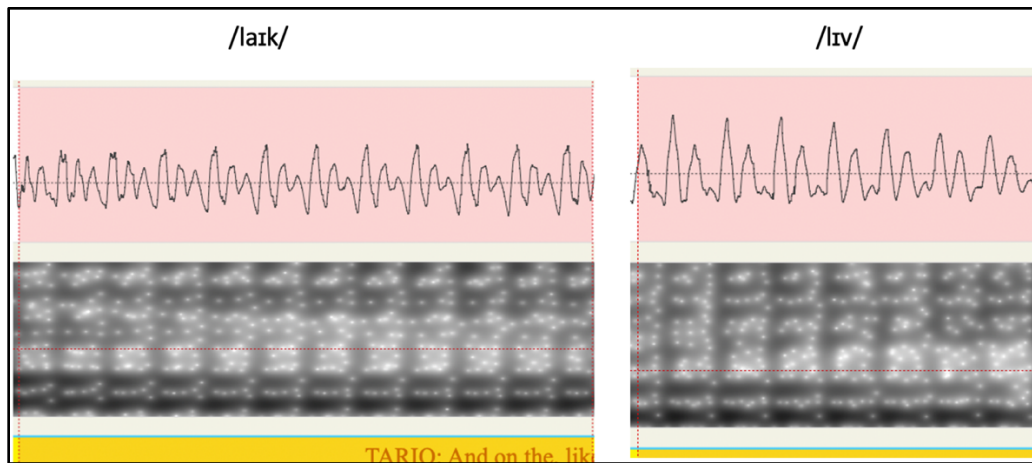
Figure 7: Spectrogram of Tariq's prevocalic /l/ production for *like* and *live*

The production of clear /l/ above shows F1 and F2 as being close to each other with a higher separation for F3. Figure 8, however, showing Tariq's production of dark /l/ *small, will, girl* and *still* point to two main spectrographic features. The first shows a greater separation between the 1st and 2nd formants and a lower 3rd formant. The second feature shows irregularities for F1 and F2 with valleys forming between these two formants (as highlighted by the yellow box). Based on these spectrographic views, a clear and dark /l/ production is possible in PakE.
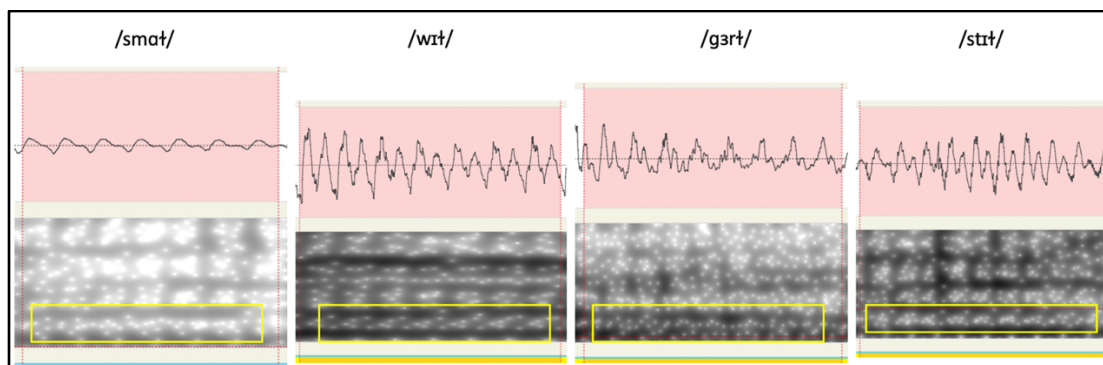

Figure 8: Postvocalic /l/ production from Tariq's speech

The final feature distinguishing PakE from GAE is the lack of aspiration in voiceless stops for word-initial /p/ and /k/. During the interview, Tariq used a total of four words that would be aspirated if spoken by an American speaker (i.e., *continuously, park, put* and *paper*). Although quite noticeable as an accentedness feature, this lack of aspiration in PakE can also have intelligibility consequences. The functional load weight for word initial /p/ or /k/ is as high as 98% and 100% respectively (Koffi, 2021:49). This means that for the cases of word-initial /p/, it is possible that these words may be confused with a similar minimal pair that starts with a voiced bilabial /b/. Specifically analyzing the context in which Tariq produces *put* and *paper*, since no direct minimal pairs are available in GAE for these words, their intelligibility wouldn't be deemed severe. However, for *park*, some listeners may confuse it with "bark" making it unintelligible.

4.0 Conclusion

From the short conversation between two married Lx English speakers, many observations have been made regarding speech intelligibility specifically and non-native accented speech as a whole. Tabinda's most noticeable segmental differences with GAE are the vowel pair [i] vs. [ɪ], voiced fricative devoicing in the coda, and rhotics. For Tabinda, many of the segmental differences of her DomE speech are consistent with most Lx speakers from L1 Spanish backgrounds. Koffi (2021:98) confirms that the vowel pair [i] and [ɪ] are the most challenging for Spanish speakers, and consonant pairs [f] vs. [v] are problematic. Specific to Tabinda is her substitution of the final [f] for [v] and the tapping of her [r]. For her husband, Tariq, the commonality with other speakers of PakE is mainly tied to the use of dental stops and the inconsistency of choosing between [v] and [w]. In spite of these specific deviations from the expected GAE segmentals, studies have shown that Lx speakers can be intelligible while heavily accented. This intelligibility can be explained by the "redundancy" (Koffi, 2021:99) of certain Lx Englishes from specific L1 backgrounds, making them more predictable.

ABOUT THE AUTHOR

Mahdi Duris is a Ph.D. Applied Linguistics and Technology student and a research assistant at Iowa State University. The broad perspective of his research interests is in pronunciation, phonetics, and phonology. Specifically, the automation of intelligibility ratings by way of the Acoustic Masking and Intelligibility (AMI) theory. Additionally, he conducts research on Mispronunciation Detection and Diagnosis (MDD) systems. He can be reached at mduris@iastate.edu.

References

Beaulieu, S., Lundell, F. F., & Bejarano, J. (2022). Interlocutors' judgment of Lx conventional expressions: An exploratory study. *Intercultural Pragmatics*, *19*(5), 597-620.

Beaulieu, S., Reinke, K., Šebková, A., & French, L. M. (2022). I hear you, I see you, I know who you are: Attitudes towards ethnolinguistically marked French Lx speech in a traditionally homogeneous city. *Journal of Second Language Pronunciation*, *8*(2), 279-303.

Boersma, P., & Weenink, D. (2007). Praat (Version 6.1.38) [Software]. Latest version available for download from www.praat.org.

Cook, V. J. (2002). Introduction: background to the L2 user. In V. J. Cook (Ed.), *Portraits of the User*. Clevedon: Multilingual Matters.

Dewaele, J.-M. (2018). Why the dichotomy 'L1 Versus LX User' is better than 'Native Versus Non-native Speaker'. *Applied Linguistics, 39*(2), 236-240. https://doi.org/10.1093/applin/amw055.

Dewaele, J. M., & Saito, K. (2022). Positive psychology can combat the pernicious native speaker ideology. *The European Educational Researcher*, *5*(2), 225-234.

Giacomino, L. (2012). Comparative analysis of vowel space of L1 Spanish speakers and general American English. *Linguistic Portfolios*, *1*(1), 9.

Hayes-Harb, R., Barrios, S., & Tripp, A. (2021). Whose input matters? The influence of socially-differentiated input sources in adult Lx phonetic learning. https://doi.org/10.31234/osf.io/t78cu.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3), 369-377.

Levis, J. M. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, *6*(3), 310-328.

Mahboob, A., Ahmar, N. H., & Schneider, E. W. (2008). Pakistani English: Phonology. *Varieties of English*, *4*, 244-258.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45* (1), 73-97.

Munro, M. J., & Derwing, T. M. (2022). Foreign accent, comprehensibility, and intelligibility, redux. *Journal: Benjamins Current Topics: The Evolution of Pronunciation Teaching and Research*, 7-32.

Koffi, E. (2021). *Relevant Acoustic Phonetics of L2 English: Focus on Intelligibility*. CRC Press.

Koffi, E., & Simmonds, A. (2018). Gradience in Coda Cluster Devoicing in a Central Minnesota English Idiolect. *Linguistic Portfolios*, *7*(1), 11.

Wells, J. C. (1982). *Accents of English: Volume 1* (Vol. 1). Cambridge University Press.

Willis, E. W., & Bradley, T. G. (2008). Contrast maintenance of taps and trills in Dominican Spanish: Data and analysis. In *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology* (pp. 87-100). Somerville, MA: Cascadilla Proceedings Project.