

Fluency as a Marker of Speaking Skills in Bangladesh: Blind Language Proficiency Evaluations

Muhaiminul Islam Neon¹

Northern University Bangladesh

Manuscript Received: 31/10/2022

Accepted: 20/10/2023

Published: 08/02/2024

Abstract

Fluency is one of the most overt signs that convey to hearers the language proficiency of the speaker. Typically, a proficient English speaker is also fluent, and so their fluency signals their mastery of the language. Additional speaking skills indicators are accuracy and range of vocabulary. This study aims to identify and explore to what extent the common people in Bangladesh – largely considered to be an EFL (English as a Foreign Language) country – value fluency when evaluating someone's language proficiency. To that end, the researcher created various types of audio clips – each with different levels of fluency, accuracy, vocabulary, and accent – and asked a demographically diverse group of Bangladeshi people – school, college, undergraduate, postgraduate students, adult males, and females – to judge the English proficiency of all the speakers behind the audio clips. Findings revealed that fluency was the most significant marker of the ratings, accent being a close second, followed by grammaticality and, last and not least, complexity. Secondly, the adult participants tended to rate holistically while the younger generation, especially university students, discretely analyzed the speech samples. This study implies that the general people of Bangladesh should focus more on language features they can improve in – in particular, fluency – rather than something they most probably cannot – adopting a so-called standard accent.

¹muhaiminulislamneon@gmail.com

Keywords: fluency; accuracy, vocabulary, accent, pronunciation, perceived language proficiency

Introduction

The thought processes of listening to a short speaking sample and then blindly assessing them in regards to the speaker's English language competence – unfolding these is the purpose of this paper. In other words, this study aims to explore the impressions Bangladeshi people have of someone's language proficiency. The present paper focuses on fluency and three other broad language aspects – accuracy, complexity, and accent – to analyze the blind ratings of carefully crafted audio clips distributed to, and subsequent interviews with, a wide demographic range of Bangladeshi people.

Background of the Study

When it comes to language proficiency, many things come into play, for a proficient language user should have a good grasp of all four language skills – reading, writing, listening, and speaking – and a good command of grammar and vocabulary (Thirakunkovit, 2018). Scholars in the past used to debate whether the aforementioned skills need to be learned separately or whether a proficient language user is automatically good at every language skill. Thirakunkovit (2018) conducted a meta-analysis and found that the latter stance is now a minority. In other words, someone can be good at speaking but subpar in writing, and someone else can possess a robust repertoire of vocabulary but not be fluent in speaking.

However, there is a misconception among the general masses regarding language proficiency, especially speaking proficiency. They still seem confused or even oblivious to the nearly unanimous established conclusion that at least some language skills are discretely learned (Gil Vera et al., 2019). As George Bernard Shaw said, “It is impossible for an Englishman to open his mouth without making some other Englishman hate him or despise him” (as cited in Suci, 2017). This phenomenon of judging someone for the way they speak can potentially lead to discrimination in the workplace, academia, and other places (Zhu & Bresnahan, 2021).

Given the fact that English has attained the status of a world language – Siemund (2018) reported the existence of at least 76 world English varieties – it appears that people still continue to judge others at face value. Especially in EFL (English as a Foreign Language) contexts where the use of English is important, such as in postcolonial countries like India where English enjoys the prestige of being an official language (Majumder et al., 2018), the matter is even more problematic. As for the context of Bangladesh, language standards are quite disappointing despite the fact that the communicative turn in language teaching happened over two decades ago, and yet the general populace has positive attitudes towards English and seeks to improve in it (Seargeant et al., 2017).

Among the factors that signal someone's spoken language proficiency, these are the most common: fluency, accent, grammar, and vocabulary (De Jong, 2018). Where there are a number of other factors, such as overall confidence and voice quality (De Jong, 2018), these are the most prominent and are the focus of this study. Fluency in particular will be highlighted, as current research on spoken language proficiency tends to prioritize fluency over accuracy (grammar and pronunciation) (Jaramillo et al., 2020). In particular, the purpose of this study is to explore the role of fluency as a sign of someone's English language proficiency in the EFL context of Bangladesh. The study also looks at other language features – namely, accuracy, vocabulary, and accent – that some may expect to act as markers of language proficiency.

Research Questions

The following are the research questions of the study:

- a) What are the speaking sub-skills that affect language proficiency evaluations in Bangladesh, and to what extent?
- b) How, if at all, do the different demographic groups of Bangladesh differ in their language proficiency evaluations?

Literature Review

This section begins with establishing some operational definitions and proceeds to review the past literature to identify the research gaps.

Concept of Fluency

Fluency is defined in various ways in the literature. In a review study, De Jong (2018) discusses several different perceptions of fluency, which can broadly be defined as overall speaking proficiency, and narrowly considered as eloquence – the ability to use language fluidly. Contrary to general belief, fluency need not be constrained to speaking, but also extended to reading (see Förster et al., 2018) and writing (see Kim et al., 2018). However, most studies, as well as this one, focus on speaking fluency.

Regarding research in the area of fluency in speech, the buzzword is “perceived fluency.” Segalowitz (as cited in Suzuki & Kormos, 2020) is usually cited as the researcher who has categorized fluency into three parts: cognitive fluency (relating to the psycholinguistic aspect), perceived fluency (relating to the listener’s perceptions), and utterance fluency (relating to the objective attempt at evaluating fluency). De Jong (2018) finds limited attempts in the literature to objectively measure fluency – that is, utterance fluency. As for perceived fluency research, such studies tend to record samples of speech from participants and then ask raters to rate – thus reflecting the subjective nature of such research. The present study also takes this methodological route.

In regards to the factors affecting fluency (and also perceived fluency), the most frequently cited ones are pauses, hesitations, slow speech rate, and repairs. Suzuki and Kormos (2020) state that perceived fluency is affected more by speed and hesitations rather than repairs. Hesitations and pauses do not always signal a lack of proficiency, as the IELTS (International English Language Testing System) rubric distinguishes between language-related fluency – when trying to recall words and phrases and string sentences together – and content-related fluency – difficulty with the topic instead of language (De Jong, 2018). The latter is usual for

everyone, including native speakers, and so should not be confused with a (major) lack of fluency. Regarding pauses, it was established in a meta-analysis that only mid-clause and end-clause pauses significantly affect perceived fluency (Suzuki et al., 2021).

Besides these disfluency features, De Jong (2018) accounts for no fewer than 11 factors, including pace, mean length of utterance, and even voice attractiveness. However, he himself cautions against using all related factors, as many of them are interrelated with other aspects besides fluency. Therefore, for the purposes and scope of this paper, fluency will specifically refer to speaking fluency, and disfluency markers will comprise only the following: pauses, hesitations, slow speech rate, and repairs.

The Complexity-Accuracy-Fluency (CAF) Framework

Typically, people who are fluent also have good language proficiency in general (De Jong, 2018). However, that is not always the case, as a fluent language user may very well succumb to grammatical inaccuracies and be unable to use complex, low-frequency vocabulary (De Jong, 2018). The IELTS rubric also acknowledges that even native speakers (band 9 users) of English may hesitate when indulging in complex discussions (De Jong, 2018). Though many other speaking sub-skills exist, these three – complexity, accuracy, and fluency (which together constitute the CAF framework) – are most often compared and contrasted in the literature. Complexity refers to the wide range of grammatical and vocabulary knowledge reflected in someone's speech, whereas accuracy is grammatical correctness.

Suzuki and Kormos (2020) report that perceived fluency is often confused with grammatical complexity and accuracy. In other words, raters tend to conflate vocabulary and grammatical competence with fluency despite clear instructions of only focusing on fluency. The authors further categorize complexity and accuracy as learners' competence and fluency as a matter of performance. However, in perceived fluency research, we are only privy to someone's performance, so it can be argued that the entire CAF framework relates only to performance. Yet, there is credence to the

idea that fluency in particular depends largely on the individual learner's speaking style, as some studies have found high correlations with people's L1 (first language) and L2 (second language) fluency (De Jong, 2018).

Pronunciation as a Marker of Language Proficiency

Besides fluency, accuracy, and vocabulary, another factor that largely determines perceived language proficiency evaluations is pronunciation. The modern language teaching paradigm, especially in the fields of sociolinguistics and World Englishes, has moved away from the centrality of standard English, and rather vouches for a pluricentric model (Sridhar & Sridhar, 2018). This means that native-like pronunciation is no longer a goal according to language scholars. However, Isbell (2020) notes that pronunciation teaching "is back in vogue" (p. 147). For the general people in many ESL and EFL countries as well, pronunciation development is a goal, and a non-Western accent is a sigil of shame (Rani & Tina, 2020). For example, Mahjabin (2019) found that most tertiary-level Bangladeshi students seek to attain native-like pronunciation. Another Bangladeshi study by Anis and Monir (2018) even goes so far as to say, "Pronunciation is the first step to learn any language" (p. 125).

Despite the clear desire of Bangladeshi students to improve their pronunciation and reduce their accents, the results are not positive. For example, Ali (2019) conducted an empirical study on 40 undergraduates and 10 EFL teachers in Bangladesh and found that the major reasons behind pronunciation difficulties are institutional issues leading to a lack of confidence. The affective filter is so strong that in another study only 9.3% of Bangladeshi tertiary-level students admitted to having confidence in their speaking skills (Mahjabin, 2019). Furthermore, Rani and Tina (2020) conducted an inter-divisional study and found convincing evidence that rural Bangladeshis in particular struggled to learn standard English pronunciation. In fact, such pronunciation teaching research appears to be a losing battle, as second language research has shown that pronunciation is one of the only language features that become fossilized at a young age (Foote & Trofimovich, 2017). What

remains a problem, however, is that accent is a stigma and marker of perceived language proficiency.

World Englishes Research on Perceived Language Proficiency

An Asian English study on language proficiency evaluations found that ratings depend significantly on the rater's background and nationality (Edwards, 2018). In the context of Iran, Namaziandost et al. (2019) investigated the role of gender in fluency. Findings revealed that female students were better in accuracy, but male students had the edge in accuracy. In the Bangladeshi context, Ali (2019) found that more students would like to become fluent (57.5%) than accurate (42.5%). This present study can compare the findings and identify whether Bangladeshi people prefer fluency over accuracy.

Another study focused on the differences between native and non-native raters in terms of turn-taking behavior (Van Os et al., 2020). It was found that for native speakers, both premature turn-taking and delayed turn-taking affected perceived fluency negatively, but for non-native raters only the latter had a negative correlation. Another international study focused on native English speakers rating Japanese L2 learners (Suzuki & Kormos, 2020). It was revealed that native speakers tend to misconstrue comprehensibility with fluency. Indeed, it is unfortunate that the pronunciation of Japanese English varieties takes away from its fluency even in such controlled tests with explicit instructions only to evaluate fluency. Another study on Japanese English – Matsuura et al. (1999) – found that pronunciation inhibited the intelligibility to L1 raters who were not accustomed to Japanese English. However, Suzuki et al. (2021) found an unclear relation between L1 and L2 raters, and suggest further research in this area. This study fills this research gap as the rater participants are L2 users of English.

Research Gap

In short, this section has discussed the various factors related to perceived language proficiency, chief among them being fluency, accuracy, vocabulary, and accent. Few studies, to the best of the researcher's knowledge, have combinedly looked at these features,

but instead have focused on either a) one of them separately, or b) combined three of them in the CAF framework. Therefore, the first research gap this present study fills is combining these four features – fluency, accuracy, vocabulary, and accent – of oral proficiency. Secondly, perceived fluency research in the context of Bangladesh is sparse even though there is a stigma of non-standard English accents used in this country. Thirdly, a new methodology is used (elaborated in the Methodology section), but it is one based on previous studies by having raters blindly assess speech samples. Fourthly, previous language proficiency rating studies have tended to focus specifically on a single demographic, such as tertiary-level students (Ali, 2019; Rani & Tina, 2020), and so this study fills the gap by taking a cross-sectional approach by considering a diverse demography.

Finally, L2 rater participants are used in the study as per Suzuki et al. 's (2021) suggestions. Previously, there used to be a myth that the native speaker is the ideal speaker and the authority on what is grammatically correct and what is not (Sridhar & Sridhar, 2018). Robert Phillipson, Braj Kachru, and many other World Englishes scholars have vehemently attacked the prescriptive hegemony of the native speaker ideal (Sridhar & Sridhar, 2018), and so L2 raters assessing speech samples for fluency should not be problematic; rather, it fills the gap of identifying the salient features L2 users perceive when evaluating language proficiency in general, and perceived fluency in particular.

Methodology

Research Design

This study is qualitative in nature and inductive in design. It is qualitative because it tries to explore rather than prove the reasons behind the language proficiency evaluations of Bangladeshi people by focusing on fluency, accuracy, vocabulary, and accent. Moreover, the methodology of creating audio clips and asking participants to rate them is qualitative in nature – as established in the literature – unless the number of participants is substantial, at which point it would have been a mixed methods study like Namaziandost et al.

(2019) and Suzuki and Kormos (2020). However, a qualitative study alone satisfies the scope of this research study.

Participants

In order to get ratings from a demographically diverse group of Bangladeshi people, the researcher intended to collect data from school, college, undergraduate, and postgraduate students, and adults – both male and female. The participants are chosen according to the researcher’s convenience – undergraduate and postgraduate students from his circle of friends, school and college students, and adults from his relatives. In particular, two participants from each participant category were chosen, as depicted in the table below:

Table 1: Participant Profile

Partici- pants’ Pseudo- nym	Participants’ Class	Gender	Current Education Level or Occupation	Education Background
Arafat	School 1	M	Completed SSC (Class 10)	Bangla Medium
Nayan	School 2	M	Class 8 Student	English Medium
Himi	College 1	F	Studying for HSC (Class 11)	Bangla Medium
Runa	College 2	F	Studying for HSC (Class 11)	Bangla Medium
Sarin	Under- graduate Student 1	F	Final Year Undergraduate Student	Bangla Medium
Shayla	Under- graduate Student 2	F	Final Year Undergraduate Student	English Medium
Sirat	Post- graduate	F	Second Year Master’s Student	Bangla Medium

	Student 1			
Neru	Post-graduate Student 2	M	First Year Master's Student	Bangla Medium
Manik	Adult Man 1	M	Service Holder	Bangla Medium
Jellal	Adult Man 2	M	Service Holder	Bangla Medium
Keya	Adult Woman 1	F	Homemaker	Bangla Medium
Lipi	Adult Woman 2	F	Homemaker	Bangla Medium

As can be seen from the table, there are seven female participants and five male ones. All participants are referred to by pseudonyms. Regarding educational background, 10 out of the 12 participants come from a Bangla medium background. This prevalence of Bangla medium education background is also representative as most Bangladeshi people study in the Bangla medium national curriculum (Mousumi & Kusakabe, 2017). This also fits the research gap of the raters being L2 speakers. As for participant selection, purposive sampling and snowball sampling were used. Sarin, Shayla, Sirat, and Neru are the researcher's personal friends, Arafat is his cousin, and Keya and Lipi are his paternal aunts. The rest of the participants were procured through the snowball technique of asking the participants already chosen.

Data Collection Procedure

Selecting the Audio Clip Categories

Before data collection, the researcher first created various types of audio clips that are different from each other in terms of fluency, accuracy, vocabulary, and accent. The following iterations were chosen to make the audio clips:

Audio Clip 1: High fluency, high accuracy, good vocabulary, American English accent

Audio Clip 2: Medium fluency, high accuracy, good vocabulary, Bangladeshi English accent

Audio Clip 3: Low fluency, high accuracy, good vocabulary, American English accent

Audio Clip 4: High fluency, low accuracy, subpar vocabulary, American English accent

Audio Clip 5: Medium fluency, low accuracy, subpar vocabulary, Indian English accent

Audio Clip 6: Low fluency, low accuracy, subpar vocabulary, Bangladeshi English accent

The rationale behind not using the statistical method of combination (nCr) to account for every single iteration is simply because it is beyond the scope of the study. Moreover, as it is not a funded study, the researcher does not want the participants to dedicate a substantial amount of time to the rating phase. Besides, the researcher believes these six iterations are adequate to find sufficient data about at least fluency, which is the highlighted speech element of this study.

Creating the Audio Clips

The researcher at first tried to collect various audio clips that fit the chosen criteria explained above. Not finding any, he decided to create the clips on his own. To that end, he has recorded audio of himself talking about the topic, “A Rainy Day.” Six audio clips have been made adhering to the described criteria (with differing fluency and other features). Although the time duration is the same – approximately 30-40 seconds – a different script was written for each clip to accommodate different vocabulary items and grammatical competence. To combat the limitation of the voice being the same in all the audio clips, the researcher has used the software Audacity to change the audio pitch so that each audio clip has a separate voice. As for speaking in different accents, the researcher has spent

significant time studying different accents so he can master them enough to use them in the vastly controlled activity of reading aloud 30-to-40-second-long scripts.

The Rating Process

The six audio clips were distributed in different ways depending on the researcher's convenience. For seven participants – Nayan, Himi, Runa, Sarin, Shayla, Sirat, and Neru – the rating process involved the researcher sending the clips to their Facebook Messenger or email, after which they listened and wrote back some comments. Subsequent phone calls were made with those participants to expound on those comments and receive detailed rationale behind their ratings. As for the remaining participants – Arafat, Manik, Jellal, Keya, and Lipi – they were at the researcher's home to listen to the audio clips and gave their ratings and subsequent interviews face-to-face with the researcher. In such face-to-face situations, the researcher was able to play back the recording time and time again and selectively compare and contrast two or more difficult-to-rate clips to ensure accurate and well-thought-out ratings.

In particular, the raters were asked to listen to the clips, take notes, and comment on the speakers' English proficiency. To clarify, the researcher deliberately avoided influencing the raters' decision-making process by not mentioning specific language features like fluency and accuracy. This is done to ensure that the ratings are as open-ended and instinctive as possible. In particular, the rating process comprised two sections: one, to provide brief comments on each of the audio clips; and two, to rank the audio clips from first to sixth.

Data Analysis Procedure

The semi-structured interviews were not recorded, but the researcher made sure to clarify the ratings and comments made by the participants in his during-rating notes. The researcher expressly informed the raters that he was taking notes, and all the raters accommodated and waited patiently for the brief notes. Some of the notes were written in the notes verbatim – “coherence and cohesion

problem” – while some were short descriptions – like “Not so good.” In the case of Bangla comments used in this paper, the researcher translated them into English. Particularly close attention was given when the participants gave their rankings. Some interviewee quotes are also transcribed and translated into English if Bangla is used. Finally, the findings are presented thematically in the Findings section, and some selected quotes are used in the Discussion section where relevant.

Ethical Considerations

First of all, all potential participants were told exactly what they had to do in this study. Only those who provided consent were given the audio clips. Moreover, member checking is not required as participants are required by the design of the study to discuss their ratings with the researcher anyway. For purposes of anonymity, abbreviated pseudonyms are used when referring to them in the study.

Findings and Analysis

This section both presents the findings of the study and analyzes them. The following subsections are named after the major themes extracted from the notes taken during the rating process and the subsequent interviews.

Anonymous Best and Worst Ratings

Each and every single participant rated the speaker of Audio Clip 1 to have the best English language proficiency. This was expected by the researcher, as Audio Clip 1 had high fluency, good vocabulary, good grammaticality, and a so-called standard English accent – an American accent. Himi explicitly said that Audio Clip 1 is “very superior,” and Runa, who tried in vain to rank the audio clips, had no qualms in declaring that Audio Clip 1 is by far the best.

Regarding the perceived worst rated speech sample, it was Audio Clip 6, but this decision was not unanimous. Two of the twelve participants – Himi and Sarin – ranked a different audio clip in their last position, but they still had Audio Clip 6 in their second last

ranking. The reason behind such anonymous low ratings for the aforementioned audio clip is simply because it has been designed to have the worst of all words: subpar vocabulary, inaccurate grammar, low fluency, and a so-called non-standard English accent – Bangladeshi English accent.

As mentioned, the fact that Audio Clip 1 and 6 are the best and worst rated clips respectively is no surprise. In essence, these two clips served as a benchmark for raters to compare the other clips. In other words, the researcher purposefully designed a model audio clip (Audio Clip 1) which can serve as a standard, as well as an intentionally lower cap (Audio Clip 2) to put things into perspective and set a maximum and minimum range. The fact that almost all the participants were able to rate these two clips as the best and worst also signals the fact that they gave more or less thoughtful and reasoned judgments in their ratings. As a result, this finding increases the reliability of the proposed methodology.

Fluency > Accuracy > Complexity

Although findings were not anonymous or nearly so when considering the CAF framework, it was definitely the case that fluency, complexity, and accuracy all affected the ratings. Fluency in particular was a buzzword by the participants, despite the fact that the researcher did not mention the criteria of evaluation and instead facilitated open-ended and intuitive ratings. For instance, Arafat said Audio Clip 3 is accurate but not fluent, and ranked it 4th, two rankings below Audio Clip 2 which only differed from Audio Clip 3 in regards to fluency, not complexity and accuracy. Moreover, all but one rater ranked Audio Clip 6 below Audio Clip 5, and the only difference between them in the CAF framework was fluency.

After fluency, accuracy was mentioned the most. The term “grammatical mistake” was mentioned 10 times in the interviews, and the ratings reflected the fact that almost all raters took into account that grammatical inaccuracy decreases language proficiency. For instance, several raters, though not all, put Audio Clips 1, 2, and 3 above the other three. However, grammaticality was not the primary determining factor for the ratings, as no fewer than seven

raters – Arafat, Shayla, Neru, Manik, Jellal, Keya, and Lipi – ranked either one or two grammatically poor audio clips over at least one grammatically superior one.

However, it is difficult to pinpoint whether it was accuracy alone or accuracy in addition to complexity that affected the ratings. Yet, while most raters mentioned grammar in their evaluations, very few cited vocabulary. Nayan did mention the following about complexity: “The speaker clearly managed to express the feeling of getting drenched on a rainy day.” Sirat also mentioned a weakness of Audio Clip 6 to be “a wrong choice of words,” and Shayla praised Audio Clip 1 by describing it as having a “visual description.” Shayla further said that she thinks that vocabulary is “not important” and that “even with little vocab we can express ourselves.” To her and some other raters, as depicted by their ratings, complexity appeared to have the weakest correlation to high language speaking proficiency.

Power of Accent

Although the current paradigm shift of World Englishes and sociolinguistics would vouch for an inclusive and non-judgmental world where accents would not matter as long as they are intelligible, this study showed accent holds overwhelming power in regard to people’s language proficiency impressions. For example, many raters – Arafat, Sarin, Keya, Lipi – put Audio Clip 5 – which constituted medium fluency and Indian English accent – over Audio Clip 4 – which had high fluency and Bangladeshi English accent. Therefore, this study shows that Bangladeshi people tend to value a foreign English accent over a native one, even though that foreign one is also not representative of the supposed standard varieties of American or British English.

In total, variants of the words “pronunciation” and “accent” were used 19 times. This reveals the fact that almost everyone, from students to the older generation, is aware of pronunciation when evaluating language proficiency. In particular, Sirat and Runa were hyper aware of pronunciation, as they mentioned it five and six times respectively in their ratings. However, one exception is Shayla, who

felt that accent and pronunciation do not matter at all as long as they do not take away from intelligibility.

Generational and Educational Perspectives

Some patterns were found when contrasting the different rationales behind the ratings of different demographic groups in Bangladesh. First, it was found that both students and adults highly value fluency and pronunciation and, to a lesser extent, grammar. There also seemed to be more similarities among adults – especially between Keya and Lipi who gave the same ratings – than among students, who tended to vary widely in their ratings. Perhaps this is due to the many more categories of students made – school, college, undergraduate, and postgraduate students – when compared to adults – only adult males and females.

Among the students, the undergraduate and postgraduate students appeared to possess the most rational explanations and more extensive knowledge about language proficiencies. This is first reflected in the abundant comments made by them in comparison to those made by school students. As for college students, one rater – Runa – gave very detailed ratings by accounting for specific language features like grammar, accent, repetitions, and rate of speech. All four undergraduate and postgraduate students also happen to be majoring in Applied Linguistics and English Language Teaching, and due to this background, they naturally possessed keener insights into the particulars of language proficiency evaluations. Sirat's comments on Audio Clip 3 encapsulate the depth and specificity of comments offered by such raters: "Pronunciation was good and the speaker was taking pauses and stammering. However, he used fillers to fill the gap. Some grammatical mistakes but was able to correct them by himself. Overall good delivery of speech." As can be seen, very specialist knowledge of fillers and gap filling was displayed by linguistics majors. This contrasts with the older generation raters – adult males and females – whose focus was on overall language proficiency and instinctive judgments on it.

Discussion

The preceding discussion has yielded a number of intriguing results. Though there were differences in opinion among the participants, certain definite patterns emerged, one of which challenges current World Englishes scholars' opinions – as reported by Sridhar and Sridhar (2018) – of the lack of importance of pronunciation. The findings revealed that most Bangladeshi citizens not only consider pronunciation as a marker of language proficiency, but also glorify it to an excessive degree. In fact, many of them rate a less fluent speech higher than another speech having the same level of complexity and accuracy but different pronunciation. This finding is in line with Isbell (2020), Mahjabin (2019), Rani and Tina (2020), and Anis and Monir (2018) who all report that pronunciation remains a fundamentally valued element of language proficiency.

Regardless of the overwhelming influence of pronunciation, similar significance – if not more – was found to be present in fluency. In answer to the research questions, all four language variables focused on in this study serve as language proficiency indicators for both the young and old generation of Bangladesh. However, there is a marked difference in the significance of the four variables, as fluency and pronunciation appeared to be the leading factors affecting the ratings. Grammatical accuracy too was a significant factor, but complexity much less so. The following equation shows the significance level of the four factors in regard to language proficiency evaluations in Bangladesh:

Fluency $\geq$ Pronunciation $\gg$ Accuracy $\gg$ Complexity
(Note: “ $\geq$ ” refers to “greater than or equal to”; “ $\gg$ ” refers to “much greater than”)

The implications of this study are worrisome, as second language acquisition scholars have established that pronunciation is extremely difficult if not impossible to learn in adulthood (Trofimovich, 2017). More awareness should be spread so that we can tolerate native accents – if not completely abandon the hegemony of a so-called standard variety. As far as the other language variables are concerned, stress should be given more to

fluency than accuracy and complexity. In other words, impressions of someone's language proficiency depend more on someone's ability to speak in a pleasant flow rather than the speaker's grammatical correctness (which is somewhat important) and rich vocabulary (much less important). Of course, building a comprehensive vocabulary and adhering to grammar are important for students in academia, but much more than these two appears to be fluency when it comes to speaking. Therefore, this study recommends academia to focus primarily on speaking fluency, secondly on grammaticality, thirdly on enriching vocabulary, and lastly on pronunciation, if at all.

Conclusion

In summation, fluency was found to be the most significant indicator of language proficiency, and its only rival appeared to be pronunciation. While the former can be worked on, the latter is unfortunately most probably a lost cause. The other two factors – accuracy and complexity – also correlated to high language proficiency, but significantly less so. As far as the Bangladeshi demography is concerned, the older generation – adult males and females – rated similarly in terms of fluency and pronunciation, but their comments in the post-rating interview session were quite holistic. On the other hand, undergraduate and postgraduate student raters majoring in English gave extremely detailed answers by mentioning the variables – fluency, accuracy, vocabulary, and pronunciation – by name.

Limitations and Recommendations

Although the new methodological framework appears to be an overall success and has filled some research gaps, some notable limitations exist, first, in the creation of the audio clips. Granted, they were on the whole quite basic, since the researcher intentionally only used six iterations of the four variables – fluency, accuracy, complexity, and accent. Because of this limited variation in the variables – for example, only “subpar” and “good vocabulary,” and nothing in between – the researcher did not feel compelled to validate the audio clips from specialists. Moreover, it was somewhat

difficult to imitate the different accents despite the clips being approximately 30-40 seconds long. One specific difficulty came in the fact that accents are not merely discrete points but rather a continuum; in other words, people speak in a mix of accents instead of a specifically distinguishable standard variety. Therefore, for pronunciation, validation would have been wise.

A specific drawback of the execution of the methodology was seen in Audio Clip 1, as a few raters felt it was unnatural. Words like “scripted,” “reciting,” and “not a regular conversation” were used, as depicted below:

Nayan: “His accent was good and sounded like he knew what he was talking about but his fluency sounded scripted to me.”

Sarin: “Accurate fluency but the speech is delivered in a poetic style. It is not a regular conversation.”

Neru: “Good fluency and accuracy. No grammatical mistakes. It seems like you are reciting a poem.”

In hindsight, the researcher realizes his folly of overcompensating for Audio Clip 1 – in the attempt to make Audio Clip 1 the standard bearer – with fancy vocabulary and excessive complexity. Nonetheless, everyone did end up rating it the best audio clip, so there was no ultimate harm in the unnaturally fluid and complex audio clip. Still, it is recommended that future studies adopting this methodology should ensure natural spoken skills even for the standard.

Another minor drawback was observed in Audio Clip 4, which many raters felt had a high speech rate. For example, Sarin said, “The speech is very fast. It’s hard to catch the words properly,” and Neru said, “He should take a little pause when he finishes the sentences.” It was not intended for Audio Clip 4 to have a higher speech rate than the others, and unfortunately, this had an impact on the ratings. For example, Arafat, Sarin, Keya, and Lipi ranked Audio Clip 4 lower than Audio Clip 5 despite the fact that both have

accuracy and complexity, but only differed in fluency, with Audio Clip 4 being more fluent than Audio Clip 5. A final drawback came in the form of two of the raters – Arafat and Shayla – recognizing, before being told so after the rating session, that it was in fact the researcher who created the audio clips, despite the fact that the audio pitch was different for each clip. Of course, only those close to the researcher were able to discern this as they were familiar with the researcher's idiosyncratic vocal patterns despite the altered pitch.

On the whole, validation of the audio recordings appears to be wise even with complex iterations left out. Despite these clear drawbacks, the methodology is recommended to be used by future researchers in perceived language proficiency evaluations.

References

- Ali, M. M. (2019). Pronunciation Difficulties of the EFL Undergrads in Bangladesh: An Empirical Study. *Journal of ELT and Education*, 2(2), 19-25.
- Anis, M. A. R., & Monir, M. (2018, December). Teaching pronunciation to primary level Esl learners in Bangladesh: A critical pedagogy perspective. *Journal of Science and Technology*, 8(1 & 2), 125-135.
- De Jong, N. H. (2018). Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly*, 15(3), 237-254. <https://doi.org/10.1080/15434303.2018.1477780>
- Edwards, J. G. H. (2018). Sounding native: Rating speakers of English from Asia. *Asian Englishes*, 21(3), 243–263. <https://doi.org/10.1080/13488678.2018.1464093>
- Foote, J. A., & Trofimovich, P. (2017). Second language pronunciation learning: An overview of theoretical perspectives. *The Routledge handbook of contemporary English pronunciation*, 75-90. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315145006-6/second-language-pronunciation-learning-jennifer-foote-pavel-trofimovich>
- Förster, N., Kawohl, E., & Souvignier, E. (2018). Short-and long-term effects of assessment-based differentiated reading instruction in general education on reading fluency and reading

- comprehension. *Learning and Instruction*, 56, 98-109.
<https://doi.org/10.1016/j.learninstruc.2018.04.009>
- Gil Vera, V. D., Jaramillo, B., Cardona, N. B., Cifuentes, M. A., Jimenez, S. A., & Martínez, L. M. (2019). Should an Effective Language Learning Be through the Development of Just One Language Skill?. *The Qualitative Report*, 24(11), 2778-2793.
<https://doi.org/10.46743/2160-3715/2019.4015>
- Isbell, D. R. (2020). Diagnostic language assessment for L2 pronunciation: A worked example. *Pronunciation in Second Language Learning and Teaching Proceedings*, 11(1).
- Jaramillo, M. G. L., Dávila, A. M., & Jarquín, E. V. E. (2020). English as a Lingua Franca: Perspectives from International MA TESOL Students in the United States. *Universal Journal of Educational Research*, 8(4), 1605-1619.
<https://doi.org/10.13189/ujer.2020.080454>
- Kim, Y. S. G., Gatlin, B., Al Otaiba, S., & Wanzek, J. (2018). Theorization and an empirical investigation of the component-based and developmental text writing fluency construct. *Journal of Learning Disabilities*, 51(4), 320-335.
<https://doi.org/10.1177/0022219417712016>
- Mahjabin, M. (2019, July 1). Role model for non-native speakers' English pronunciation: 'an investigation into the attitude of adult learners in Bangladesh'. [Unpublished master's thesis].
<http://dspace.bracu.ac.bd/xmlui/handle/10361/13795>.
- Majumder, S. D., Verma, N., Sonkar, S., & Kumar, S. (2018). Language policy for assigning English as the official language and promoting regional language as a compulsory subject in academics to preserve a country's ethnicity. *International Journal of English Learning & Teaching Skills*, 1(1), 66-70. <https://doi.org/10.15864/ijelts.1111>
- Matsuura, H., Chiba, R., & Fujieda, M. (1999). Intelligibility and comprehensibility of American and Irish Englishes in Japan. *World Englishes*, 18(1), 49-62. <https://doi.org/10.1111/1467-971X.00121>
- Mousumi, M. A., & Kusakabe, T. (2017). The dynamics of supply and demand chain of English-medium schools in Bangladesh. *Globalisation, Societies and Education*, 15(5), 679-693.
<https://doi.org/10.1080/14767724.2016.1223537>

- Namaziandost, E., Abedi, P., & Nasri, M. (2019). The Role of Gender in the Accuracy and Fluency of Iranian Upper-intermediate EFL Learners' L2 Oral Productions. *Journal of Applied Linguistics and Language Research*, 6(3), 110-123.
- Rani, S., & Tina, A. A. (2020). The Impact of Bangla Regional Dialect on the Pronunciation of English at Tertiary Level. *Humanities & Social Sciences Reviews*, eISSN, 2395-6518.
<https://ssrn.com/abstract=3573592>
- Seargeant, P., Erling, E. J., Solly, M., Chowdhury, Q. H., & Rahman, S. (2017). Analysing perceptions of English in rural Bangladesh. *World Englishes*, 36(4), 631-644. <https://doi.org/10.1111/weng.12215>
- Siemund, P. (2018). Modeling World Englishes from a cross-linguistic perspective. *Varieties of English Around the World*, 133–162.
<https://doi.org/10.1075/veaw.g61.06sie>
- Sridhar, S. N., & Sridhar, K. K. (2018). Coda 2 A bridge half-built: Toward a holistic theory of Second Language Acquisition and world Englishes. *World Englishes*, 37(1), 127-139.
<https://doi.org/10.1111/weng.12308>
- Suciu, G. (2017). What's in an accent?. *Journal of Romanian Literary Studies*, (12), 518-521.
- Suzuki, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, 42(1), 143-167.
<https://doi.org/10.1017/S0272263119000421>
- Suzuki, S., Kormos, J., & Uchihara, T. (2021). The Relationship Between Utterance and Perceived Fluency: A Meta-Analysis of Correlational Studies. *The Modern Language Journal*, 105(2), 435-462.
<https://doi.org/10.1111/modl.12706>
- Thirakunkovit, S. (2018). The Notions of Language Proficiency and Language Dimensionality. *PASAA: Journal of Language Teaching and Learning in Thailand*, 55, 219-236.
<https://eric.ed.gov/?id=EJ1191745>
- Van Os, M., De Jong, N. H., & Bosker, H. R. (2020). Fluency in dialogue: Turn-taking behavior shapes perceived fluency in native and

nonnative speech. *Language Learning*, 70(4), 1183-1217.
<https://doi.org/10.1111/lang.12416>

Zhu, Y., & Bresnahan, M. J. (2021). A thematic analysis of international teaching assistants' stigma experience in a US university: English-proficiency determinism. *Journal of International and Intercultural Communication*, 14(2), 146-163.
<https://doi.org/10.1080/17513057.2020.1762110>