Investigating the Policy-Reality (Mis)Match in IELTS and TOEFL from the Perspectives of Global Englishes

Yusop Boonsuk - Ali Karakaş*

Received: November 2, 2019; received in revised form: December 22, 2019; accepted: January 10, 2020

Abstract:
Introduction: In recent years, the number of test-takers of international tests of English has grown at an exponential rate. Those whose first language is not English, i.e. non-native English speakers (NNES), constitute the predominant majority of these test-takers, largely based in non-Anglophone contexts. Thus, the state of whether the international tests of English are fit for purpose and reflect the current realities of language users, especially in higher education institutions, has become a matter of serious concern recently. Hence, we aim to analyze the websites and language documents of two major international tests of English boards (i.e. IELTS and TOEFL) in relation to the kind of Englishes against which they judge their test-takers’ English proficiency - either by implication or by explicit expression.

Methods: To analyse the websites and language policy documents of the major international tests of English boards, we adopted a qualitative research design in which our prime purpose was to collect a blend of textual, visual and audio materials from their websites as well as publicly available documents, such as skill-band-descriptors, sample test materials, and handbooks for test takers. The analysis of the data was multimodal, utilizing a mixture of qualitative frameworks to analyze the websites and documents.

Results: The findings reveal that IELTS and TOEFL promote themselves as welcoming international test takers, while in practice, most of the contents in their examinations still draw on NES norms based on what is considered standardized English. Visual portrayals on their websites indicate that these tests are aware of English diversity and aim to embrace multicultural clients. However, no remedial measures seem to have been taken in practice as can be understood from their test and measurement criteria as regards writing and speaking.

Discussion: Drawing on the results, it may be argued that the visual portrayals are merely the tools to attract NNES test takers and the covert message is that those who are NNESs should take these proficiency

* Yusop Boonsuk, Prince of Songkla University, Songkhla, Thailand; yusop.b@psu.ac.th
Ali Karakaş, Burdur Mehmet Akif Ersoy University, Burdur, Turkey; akarakas@mehmetakif.edu.tr
examinations. Moreover, many listening exams employ NESs to produce the voices or simulate the conversations. Although the tests claim that the voices are from diverse accents, including Australia, Canada, New Zealand, the UK, and the USA; they fail to recognize that many more English varieties exist within the Outer and Expanding Circle countries.

Limitations: This research has only dealt with two major international tests of English, namely IELTS and TOEFL. There are other major tests of English available in the market. Therefore, sufficient caution should be exercised while generalizing the results to other tests as there may have been some rethinking and awareness in other tests with respect to their future test-takers’ profile and linguistic diversity.

Conclusion: The findings illustrate a degree of recognition of Global Englishes (WE and ELF) at a “theoretical level” in the international tests of English, but at the “practical level”, many crucial principles are absent as the tests, judging international test takers, remain confined within the native-norm territory. In short, the phenomenon demonstrates a theoretical level of awareness, but such awareness is not further applied at the practical level.

Key words: global Englishes, world Englishes, English as a lingua franca, assessment, (non)native English speakers, standard English.

Introduction
During the past decade, the number of individuals taking standardized English proficiency tests has risen steeply. Most of these test-takers are from countries where English is not spoken as the first language. To survive in the globalized world in which “language and intercultural competences have become core employability skills in many fields” (Luka & Seniut, 2019, p. 1), many non-native English speakers (NNES) resort to taking English proficiency tests, designed by renowned international examination boards, such as TOEFL and IELTS, to provide clear proof to education institutions, particularly those of higher education, and employers that their English qualifications are adequate. The fact that most test-takers are non-native is a reflection to a larger sociolinguistic phenomenon, which is that most English speakers are also the non-natives who reside in the Expanding and the Outer Circles (Kachru, 1986) and use English as a lingua franca (ELF). Consequently, the roles of English have shifted in the grand scale and more global Englishes have merged. Considering the test-taking demographic and the use of ELF, implementing standardized English proficiency tests on non-native test-takers has raised a suitability controversy which is whether it is justified to primarily measure non-native test-takers with the native linguistic norms. While most non-native test-takers mainly use English as a lingua franca in cross-cultural interactions with minimal number of native English interlocutors, “the legitimacy of using standardized tests, which take Inner Circle varieties of English as the norm for
local institutions, has been highly criticized in recent years.” (Arik & Arik, 2014, p. 8) Contrary to Lowenberg’s (1993) assertion that global-scale measurements of English proficiency should be standardized using common English norms approved by highly educated native English speakers, Jenkins (2014) argues that testing institutions should not employ British or American English conventions in assessing international test-takers’ English whose English-application purposes are tied to English diversity and intercultural communication. Moreover, English tests, such as IELTS and TOEFL, are too broad to be a reliable indicator and too insensitive to forecast future academic outcomes for specific English user groups (Leung, Lewkowicz, & Jenkins, 2016). Backing the argument, Davies, Hamp-Lyons, and Kemp (2003) noted that these tests neglect English diversity, discriminate speakers of other English varieties, and force all English users to adopt English nativeness through score deduction and penalization. To understand this test-taking phenomenon better, this study aims to analyze websites and language documents of major international tests boards of English proficiency to reveal the English varieties used to measure the test-takers; compare the measurements with the explicit marketing expressions; and confirm whether the claims favor any specific English variety over the others.

1 Global Englishes: World Englishes and ELF

Global Englishes is a sociolinguistic paradigm broadly invented to accommodate linguistic and sociocultural issues which occurred among global English users (Centre for Global Englishes, 2018). While the term is aimed to minimize geographic, linguistic boundaries and differentiation of language varieties, it is also expected to enhance the importance of English pluricentricity and fluidity (Galloway & Rose, 2015). Many studies view “Global Englishes” as a term for a larger conceptual classification containing World Englishes (WE) and ELF ideologies (e.g. Galloway & Rose, 2015; Jenkins, 2011; Jenkins, Cogo, & Dewey, 2011; Widdowson, 2015). While WE deals with “the identification and codification of national varieties of English”, ELF “examines English use within and across such borders, as well as focusing on the global consequences of English’s use as a world language.” (Galloway & Rose, 2015, p. xii) Despite the similarities they share, both ideologies deal with different aspects of English pluricentricity. Galloway and Rose (2015) consider both terms similar as they share a widespread pluricentricity, non-native focus for English interactions, liberation of English ownership, and transformation for English language teaching.

However, WE differs from ELF as WE emphasizes the codification of national varieties which is mainly in the Outer circle of Kachru’s (1986, 1992) model whereas, ELF emphasizes the uses of English in all the three circles: Inner, Outer, and Expanding with the Expanding Circle being its special interest. Nonetheless, as ELF is about linguistic accommodation where English norms can be linguistically negotiated and reshaped by the users according to
situational demands, it does not seek to codify emergent English use among speakers (Galloway & Rose, 2015). Based on the abovementioned differences and similarities, we produced the following figure to highlight the overlaps and divergences between the research paradigms.

Figure 1. Differences and similarities between WE and ELF.

Combining the similarities and differences between WE and ELF, the collective concept of Global Englishes comprises the following ideological characteristics: 1) most English speakers are non-native and they should not be perceived as “eternal learners” or “failed native speakers” but “successful communicators”; 2) linguistic diversity and divergence should not be viewed as problematic in communication; 3) native English speakers should not be considered an ideal model or target for language learning; and 4) English, currently, belongs to the global community, and native English speakers should no longer claim the ownership (Galloway, 2013; Galloway & Czarnecki, 2014; Galloway & Rose, 2015; Jenkins, 2011; Jenkins et al., 2011).

2 Global Englishes: English language assessment

Since the current ELT practices do not conform to the discussed ideological characteristics of Global Englishes, it is essential to reconsider the teaching direction. In an ELF perspective, English users should receive a performance assessment that is ELF oriented, e.g. accounting ELF awareness and factors. Specifically, standardized English tests neglect the fact that English is a pluricentric language and employ nativeness; the monocentric English features, such as phonological and syntactic norms, to measure how well an ELF user uses the pluricentric language. Consistently, Davies et al. (2003, p. 571) indicate that these standard tests take “an International English (IE) view [which] insists that the only acceptable norms are those of native English speakers (NES).” To
solve the issue, Hu (2012) determines the linguistic testing norms and the choice of English variety based on the intended use and social preferences. Although many English encounters would involve NESs, many more occasions would involve other interlocutors with different English varieties. Hence, an ELF- and WE-friendly test should measure test-takers based on intercultural and communicative competence. As Canagarajah (2007) insists, such tests should determine how well a test-taker can understand communicative intents, behaviors, and expectations that other English speakers from different linguistic and cultural backgrounds attempt to convey. In his words, test-takers should be scored based on their “strategies of negotiation, situated performance, communicative repertoire and language awareness” rather than grammatical accuracy (p. 936). Furthermore, Jenkins (2007) also agrees that accommodation skills should be prioritized in ELF proficiency tests and any linguistic divergence commonly produced EFL and ELF users should not be viewed as mistaking. In short, such users should not be penalized as long as the divergent language is intelligible and does not lead to a communication breakdown. Consistently, Khan (2009) supports that an English proficiency assessment should be adaptive to suit all the requirements of intercultural communication as these interactions occur among members of global communities whose language backgrounds are not always mutual.

Likewise, replacing the measurements for English nativeness, Harding (2012) maintains that the following determining factors should be the new focuses of English competence: 1) English varieties: accents, syntactic forms, and discourse styles; 2) meaning negotiation; 3) phonology for meaningful intercultural communication; 4) intercultural pragmatics; 5) target audience and communicative priority and; 6) strategic competence to avoid communication breakdowns. Practically, forcing a paradigm shift in the test arena could take time as WE and ELF measurements are in the ongoing process of development. Matsuda (2012) believes that the concepts are still new to teachers and not many are equipped to teach their learners every English variety. Depending on limited learning resources on English diversity, learners will still struggle. Since most of the formal English learning materials are distributed from a few British and American firms (Dewey, 2015; Jenkins, 2007; McKay, 2002; Methitham, 2009), many WE and ELF educational stakeholders will continue to feel influenced by a false value implicating that proficient English users are the ones who know native Englishes well whereas the incompetent ones are those who do not (Jindapitak & Teo, 2012).
3 Research methodology
This study employed a qualitative approach to examine the major international examination boards, TOEFL and IELTS, through their websites and test policies. The data were collected from diverse and publicly available sources of textual, audio, and visual data, especially those indicating skills, bands, test samples, and testing guidelines. In general, the investigation looked for policy and examination materials that could indicate some links to specific English-test designs, directions, and decisions (Blaxter, Hughes, & Tight, 2010). From many pages of the TOEFL and IELTS websites, textual (e.g. assessment criteria and rubrics, test formats, score calculation, teachers’ guides, practices, and test samples); visual (e.g. advertised images and videos of potential test-takers and test-taking experiences); and audio data (e.g. listening test-samples) were collected. The data were then analyzed to determine these examination boards’ accommodations for diverse test-takers with different English backgrounds; uses of specific English varieties in media and assessment; and pressure on users of diverse English backgrounds towards a few English varieties deemed standard. Knox’s (2007) three-dimension model, i.e. Macro-Theme-Rheme, Head-Tail, and Primary-Secondary was employed for the visual (website) analysis. Knox’s (2007) groupings are useful in identifying the degree of importance attached to the visual materials on the websites. For example, according to Johnson, Milani and Upton (2010, p. 231), on websites, the primary emphasis is placed on “the left-top corner of the screen immediately below the browser window” as it is “the guaranteed viewing area on any computer screen and is therefore a strategic location” for visitors. The Macro Theme-Rheme dimension deals with this location in the main. With reference to the head-tail dimension, it requires visitors to navigate through a web page on a certain website by scrolling up and down through text or graphics on the website. The point behind this is that “the hierarchy of information is typically indexed according to a continuum of diminishing salience.” (p. 230) That is, a visitor can easily check over imperative information in a comparatively short time as they appear in areas straightforwardly accessible whilst less important information is dispersed towards the bottom of the page, thereby reachable by visitors in a longer period of time. The final dimension, Primary-Secondary, relates to whether information is positioned in the middle of the webpage or the right- or left-hand columns of it. As Knox (2007) notes, the middle column is apportioned to important information whereas the left- and right-hand columns are subordinate in terms of information they circulate.
To analyze the textual (documentary) data, qualitative content analysis (Schreier, 2012) and negative analysis were employed. In our analytical framework, we followed a four-step process: “finding, selecting, appraising (making sense of), and synthesising data contained in documents.” (Bowen, 2009, p. 28). While engaged in this process, we were interested in “the deep structural meaning conveyed by the message” rather than the surface level meaning of the message (Berg, 2001, p. 242). We consolidated the qualitative content analysis with “negative analysis”, particularly paying mind to “meaningfully absent” mentions as regards to some issues of language use and assessment in the documents (Pauwels, 2012, p. 253). Additionally, to ensure validity, thorough extracts from related policy statements and documents, together with accessible links to the relevant pages were given in the footnotes so as to show that our analysis is “solid”, “comprehensive” and is “presented in a transparent way, allowing readers, as far as possible, to test the claims [and interpretations] made.” (Jørgensen & Phillips, 2002, p. 173)

4 Findings and discussion

4.1 Multimodal website analysis
The first step of the website analysis was to examine the visual data. The primarily derived visual data were images of people displayed at the homepages of both examination boards. The purpose of this visual investigation was to identify which of Kachru’s (1986) concentric circles are the origins of these displayed people. The findings revealed that the people displayed on these websites’ homepages tend to be from Outer- and Expanding-Circles. The images were strategically placed immediately at the beginning of the homepages to catch attention. The images were placed to ensure to be the very first elements the websites’ visitors will digest. More importantly, none of the images represented an Inner-Circle speaker. Specifically, the images primarily featured East and South Asian individuals presumably to indicate the major groups of
test-takers in the world. Featuring non-native English speakers (NNESs) on the examination boards’ webpages implicates the awareness of the dominating sociolinguistic profiles of English speakers which are those from the Outer and Expanding Circles speaking English as a non-first language.

The second step involves negative analysis. Featuring no native English speakers (NESs) or Inner-Circle speakers could be interpreted that these examination boards mainly aimed to communicate with NNESs as they are the group that is desirous to measure English proficiency and use these proofs to acquire various professional and personal gains, e.g. education, work, and migration. As most Inner-Circle speakers do not need to take these tests to prove their English proficiency, the major applicant group of these examinations remains primarily non-native. Consequently, the websites of these international examination boards incorporate images of non-Inner-Circle citizens conveying that the tests are opened to them as prospective test-takers.

![Figure 3. An image displayed on the IELTS homepage.](https://www.ielts.org/)

![Figure 4. An image displayed on the TOEFL homepage.](https://www.ielts.org/)

The third step was an in-depth investigation of the visual and textual data in Head-Tail and MacroTheme-Rheme areas. Both homepages were browsed by scrolling down, left, and right to inspect visual and textual elements. After scrolling down the IELTS homepage, the web banner immediately appeared right below using the image of Figure 1 describing that “IELTS is the high-stakes English test for study, migration or work.” (What is IELTS\(^1\), 2019) Following the banner towards the left column was a video, “Filip’s IELTS story” which seems to have been given some priority based on Knox’s (2007) MacroTheme-Rheme analysis. The video is 1:31 minutes long depicting some scenes in London and University College London as Filip was accepted for the university’s admission for passing IELTS. It is noteworthy to mention that, in the video, Filip spoke English with his Bulgarian accent depicting that he was a successful non-native test-taker and that he could speak English his way without being forced to adopt pronunciations and accents of standard Englishes.

In the final step, areas at the end of the homepages were investigated. For the IELTS homepage, five out of six images displayed the Inner-Circle landscapes taken in countries, such as Australia, Canada, New Zealand, UK, and USA, in

\(^1\) https://www.ielts.org/
response to the bold red caption, “Travel the world with IELTS” whereas the remaining image displayed a world map with country dots indicating that the test is globally recognized and accepted by governments, organizations, and universities paired to a question, “Where can IELTS take you?” as demonstrated in Figure 3. From the Primary-Secondary and Head-Tail concepts, the images of the top destinations are positioned in all the prime positions whereas the world-map one was located further to the right and way bottom. This could be interpreted as the world map was perceived as less significant than those top destinations in the eye of IELTS. Another assumption was that IELTS believes test-takers take its test because they want to enter an Inner Circle destination. Nonetheless, recent studies spoke otherwise. More test-takers sit for an IELTS examination to acquire a proof of English proficiency for academic purposes, e.g. to satisfy an admission requirement of a university. As many universities that employ English as a Medium of Instruction (EMI) require such proof, prospective students are forced to take IELTS. These universities are not only situated in the mentioned Inner-Circle destinations. Conversely, many of them are within the countries where the non-native test-takers reside. As a result, prioritizing displays of Inner-Circle destinations on the IELTS homepage indicated examination board’s misconception of the new reality. It was why the board decided to position itself as a road to Inner Circle countries.

Figure 5. First screen of IELTS home page.

Figure 6. TOEFL homepage.
Moving on to the TOEFL homepage and basing the observation on the MacroTheme-Rheme perspective, the examination board uses the left side of the page column to display figures of institutions which utilize TOEFL scores. Quite consistently with the IELTS homepage, TOEFL firstly states names of Inner Circle countries, secondly mentions Europe and Asia, and lastly reports countries recognizing the scores. Towards the left in a promotional message, the examination board states:

TOEFL scores are accepted by more than 10,000 universities and other institutions in over 150 countries, including Australia, Canada, New Zealand, the U.K., the United States, and all across Europe and Asia (The TOEFL Test, 2019).

The statement indicates that TOEFL is aware that most of its test-takers would be using their scores in countries that are not English-speaking, i.e. in Europe and Asia. Despite being aware of the countries accepting TOEFL scores, the examination board’s homepage was primarily about some Inner-Circle countries whereas others are mentioned collectively in continents. This action could be viewed as the board’s attempt to communicate to its audience that TOEFL is a gateway to the Inner Circle. The left column of the page contains three rows which are arranged vertically from top to bottom. The top row contains a short TOEFL story of Samir with a link to a more detailed one on a separate page. This detail page comprises a video at its central area illustrating scenes of Samir’s life from his childhood through his adulthood. One of the scenes shows Samir holding a TOEFL preparation book during his teenage days, and another scene shows his later days saying “Thank You” on his graduation day at some English-speaking university. According to the way in which he sounds, Samir is an English speaker who seems to have some Indian background.

Right below Samir’s is “TOEFL MOOC” and a video link section introducing “The Insider’s Guide to the TOEFL Test Course.” The video section begins with images of NNESs, a native English speaker giving a brief course description, and some test designers talking about the creations of TOEFL tests.

The rationale of why TOEFL opts to showcase images of test-takers from the Outer and Expanding Circles could be because most of its clients are genuinely from these non-Inner-Circle regions who wish to be tested by Inner-Circle language experts and acquire their TOEFL scores to prove the language proficiency levels. To support the notion that these clients need to be tested by native language experts, some concrete evidence is available towards the end of the TOEFL test preparation webpage. In this page, most of the instructors are

---

2 https://www.ets.org/toefl
3 https://www.edx.org/course/toeflr-test-preparation-insiders-guide-etsx-toeflx-4
4 https://www.edx.org/course/toeflr-test-preparation-insiders-guide-etsx-toeflx-4
NESs except one, i.e. the listening instructor who looks rather Chinese. Moreover, the publicly available listening samples are NES-produced\(^5\).

4.2 Documentary data analysis

The document data, especially those on speaking, writing and listening, derived from the examination boards were analyzed to determine the following questions: 1) whether both of these tests have been mainly developed or their test-takers are being judged based on a specific English variety; 2) whether there has been any accommodation of WE and ELF (as suggested by Harding, 2012; Hu, 2012; Jenkins, 2007; McNamara, 2012; Widdowson, 2012) in the tests; and 3) whether there has been a recognition of WE and ELF along with traces of deviant and non-normative English in the tests.

Initially, the documents on speaking were investigated. In relations to speaking, the analysis looked for a scoring mechanism, assessment, and traces of implicit or explicit requirements with a possible tie with a specific variety of English. To answer the questions, content analysis was employed on these documents, e.g. speaking rubrics, speaking test sample, and test-taker handbooks. In general, the examination boards employ matrices, e.g. fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation in the tests.

4.2.1 Speaking assessment and scoring

Beginning with IELTS, the “IELTS Speaking - How it’s marked\(^6\)” section indicates that test-takers must be able to talk using normal continuity, rate, and effort; and use ideas and language to create coherent and connected speeches through key indicators of fluency, speech rate, and speech continuity.” (“IELTS test format”, 2019) For grammatical range and accuracy, IELTS requires the grammatical range to cover spoken sentences with proper length and complexity, appropriate subordinate clauses, and adequate knowledge of sentence structures to be able to switch around elements for different communicative emphases with the key indicators being “the number of grammatical errors in a given amount of speech and the communicative effect of error.” (“IELTS test format”, 2019) In terms of pronunciation, speakers are judged on the basis of “the amount of strain caused to the listener, the amount of the speech which is unintelligible and the noticeability of L1 influence.” (“IELTS test format,” 2019) Clearly, some substantial criteria are not provided as to how normal is deemed reasonable, how continued the talking continuity should be, which variety of English is the correct English, and why traces of first languages in speech could be determined as flawed or mistaking.

As for TOEFL, its “Independent Speaking Rubrics” assess speakers using a 4-point rating scale covering delivery, language use, topic development, and

\(^5\) https://www.examenglish.com/TOEFL/toefl_listening.htm

\(^6\) https://www.ielts.org/about-the-test/test-format
general description. To obtain 4 points, which is the highest score of the speaking section, the following criteria must be met:

- **Delivery**: Generally well-paced flow (fluid expression). Speech is clear. It may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility.

- **Language use**: The response demonstrates effective use of grammar and vocabulary. It exhibits a fairly high degree of automaticity with good control of basic and complex structures (as appropriate). Some minor (or systematic) errors are noticeable but do not obscure meaning (our italics).

Apparently, TOEFL values intelligibility and is willing to tolerate minor errors as long as such mistakes do not alter meanings in communication. Nonetheless, it also expects “effective use of grammar” and “a fairly high degree of automaticity.” When looking subjectively into the descriptions, many descriptors could be controversial. For instance, when TOEFL expects production of clear speeches, clarity can be the controversy as listeners can have different opinions about how clear is clear. Specifically, speeches that a non-native speaker produces could be comprehensible to NNES- but unclear to NES-listeners. The decision to use clarity without defining subjectivity indicates that TOEFL neglects to take test-takers’ sociolinguistic backgrounds into consideration. Although it is willing not to penalize minor errors, as previously discussed, it fails to recognize that, in real communicative encounters among the non-native audience, ELF speakers often employ communication strategies to facilitate their limited English, e.g. shortening sentences and using long pauses.

Evidently, TOEFL will penalize test-takers and give them 1 out of 4 scores when they fail to provide “consistent pronunciation stress and intonation difficulties” and “cause considerable listener effort.” (TOEFL Speaking Rubrics⁷, 2014, p. 1) This means that when a test-taker deviates from standard stresses and intonations or give listeners a hard time trying to understand the communication, such a speaker will get 1 score as a penalty. Nonetheless, since communication is a two-way effort requiring both parties (i.e. speakers and listeners) to try to achieve a mutual understanding, listeners should also be responsible for ensuring that the communication is carried out effectively. Especially in ELF contexts, conversations will not always be smooth so both parties should take an equal share of responsibility to facilitate the communication.

### 4.2.2 Writing assessment and scoring

According to the IELTS test format⁸, test-takers’ writing skills are assessed and scored based on their completion of tasks involving a short description of a graph or chart and a long academic passage on a designated topic. The

---

⁷ [https://www.ets.org/s/toefl/pdf/toefl_speaking_rubrics.pdf](https://www.ets.org/s/toefl/pdf/toefl_speaking_rubrics.pdf)

⁸ [Academic Writing at https://www.ielts.org/about-the-test/test-format](https://www.ielts.org/about-the-test/test-format)
completeness is measured through the following criteria: task response, coherence and cohesion, lexical resource, and grammatical range and accuracy. Dimensionally, lexical resource requires test-takers to possess a “range of vocabulary used and its accuracy and appropriacy in terms of the specific task” (“IELTS test format”, 2019) whereas grammatical range and accuracy requires a “range and accurate use of grammar as manifested in their sentence writing.” However, what IELTS expects for coherence and cohesion is rather unanticipated: “overall clarity and fluency of the message.” Apparently, this test matrix requires a level of conformity with standard English norms where grammatical correctness and accuracy is paramount. In other words, writers with knowledge of academic writing conventions in standard English will most likely be the ones earning high scores in coherence and cohesion and lexical resource. Similar notions are stated in the IELTS guide for teachers⁹ indicating that test-takers achieving Band 9, described as Expert English Users, are those who can use “a wide range of structures with full flexibility and accuracy; rare minor errors occur only as ‘slips’.” (IELTS guide for teachers, 2017, p. 23) This means that errors, deemed as slips, might be ignored but the use of non-standard English, such as WE and ELF, in writing could hurt the assessment outcomes and scores.

In the TOEFL writing examination, “Independent Writing Rubrics¹⁰” were among the first analyzed. These rubrics determine test-takers writing proficiency and score them on a 5-point rating scale: from 0 to 5. Evidently, the scores are measured through the criteria on grammatical correctness using native lexical norms. To get 5 point which is the highest score, test-takers must demonstrate “consistent facility in the use of language, demonstrating syntactic variety, appropriate word choice, and idiomaticity, though it may have minor lexical or grammatical errors.” (TOEFL Writing Rubrics, 2014, p. 1) Not only the writing test requires the clients to demonstrate their repertoires of native English norms through writing; it also tells the world that Standard English is exclusive to any of the native-English varieties. Since NESs are more comfortable with idiomaticity, most NNESs might not always use them appropriately. Often, NNESs would either avoid, find a workaround, or reinvent new ways of saying things rather than employing the existing idioms, and these solutions are perceived incorrect in testing.

4.2.3 Speaker profiles in listening exams
In the analysis of listening-related documents, speaker profile was, instead, the prime analytical focus, and as a result, listening criteria and assessment were not of the essence. Specifically, the analysis aimed to inspect the origins of the

⁹ https://www.ielts.org/-/media/publications/guide-for-teachers/ielts-guide-for-teachers-2017-uk.ashx
¹⁰ https://www.ets.org/s/toefl/pdf/toefl_writing_rubrics.pdf
voices implemented in the listening-test materials. Initially, relevant webpages were inspected, and vital data shared with the test-takers were gathered. After the inspection, TOEFL iBT Test Content\textsuperscript{11} revealed the following information on listening and speaking:

- Native-speaker English Accents - The Listening and Speaking sections of the TOEFL iBT test include other native-speaker English accents in addition to accents from North America. You may hear accents from the U.K., New Zealand or Australia. ETS added these accents to better reflect the variety of native English accents you may encounter while studying abroad (“TOEFL iBT: Test Content”, 2019).

The provided information also confirms the notion that TOEFL is not open to other English varieties from Outer- and Expanding-Circle countries. This means that TOEFL only accepts a few English varieties within the Inner Circle. The findings support the assumption that test-takers are required to demonstrate native-English capacity within the contexts of English as a Native Language (ENL). Nonetheless, such discriminative pressure is implemented with a weak association to the current sociolinguistic reality where most English communicative encounters are produced among NNESs in non-native contexts.

Consistently, IELTS’s “Listening Test Format” plays “four recordings of NESs” to its test-takers (“IELTS Test format”, 2019) whereas its “Paper Format” states that these recordings, produced with various English accents (e.g. British, Australian, New Zealand, American, and Canadian), will only be played once during the test. The statements indicate that IELTS does not accommodate ELF test-takers in two extents. A) a one-time playback is harsh and unrealistic when in real-world communication, ELF interlocutors help each other implement many communication strategies including repetition, self-repair, and other repair to ensure the communication is smooth. B) Despite the fact that these recordings are produced in many contexts (e.g. monologue, dialogue, group discussions, and educational speeches) (“Test Format”, 2019), IELTS makes sure that these recordings are natively made.

**Conclusion**

English is now widely used as an intercultural tool for communication. Its expansion as a lingua franca introduces many new English varieties collectively called as world Englishes (WE). English is spoken mostly by NNESs who have become the dominant group of test-takers of English proficiency examinations (e.g. TOEFL and IELTS). The significant increase in number varies positively with educational and professional requirements. However, testing NNESs using native Englishes has raised a legitimacy concern as the action could discriminate NNESs who prefer different Englishes by being tested using strict NES norms. To reflect the reality of English becoming a lingua franca, English proficiency

\textsuperscript{11} \url{https://www.ets.org/toefl/ibt/about/content/}
tests should acknowledge that English is a pluricentric language and implement less NES, more performance-based, and more WE- and ELF-friendly measurement items. In this study, the producers of the examined standardized proficiency tests market that international test-takers are a key target group due to significant growth in number. For instance, their homepages show evidence that the communication is meant for the Outer and Expanding Circle audience because non-native models are used throughout the contents to convey a friendly message to the non-native group of English speakers that the tests are designed to suit them. However, the examined tests practically speak otherwise as most test items are dominantly NES-oriented. Nativeness is employed as a test matrix as it is perceived to be the standard of English. In short, these tests acknowledge the sociolinguistic profiles of their clients who are largely from a non-inner circle. Yet, the tests fail to take such sociolinguistic profiles seriously and assess them using strict NES grammar, vocabulary, and idioms. The tests neglect to measure linguistic accommodation, meaning negotiation, or intercultural communication performance. Furthermore, NNESs are found in the displayed testimonials and success stories whereas NESs (e.g. British and American) are found in the introductory videos. This signifies that the tests understand English diversity and seek to accommodate multicultural clients. Nonetheless, the tests show no sign of such accommodation in practice as NES norms dominate the test contents. The phenomenon implicates that the visual presentations are their marketing tools to attract non-native test-takers suggesting that the proficiency tests are suitable for NNESs. However, the suitability claim remains unjustified as many listening exams contain native-English voices. Many tests claim that the voices are from diversified sources of accents including Australia, Canada, New Zealand, the UK, and the USA. Nevertheless, the claim signifies a severe misunderstanding that these accents represent all the Englishes of the world when there are many more English varieties from the Outer and Expanding Circles. Hence, these standardized tests are wrong about their claims over the coverage of English diversity as their test contents remain limited to a few Inner-Circle Englishes. In conclusion, this study discovered some theoretical recognition of WE and ELF in the standardized English proficiency tests, but the practical implementation remained weak, if not missing. International test-takers’ proficiencies are still measured using native norms. The phenomenon can be interpreted that these tests are only trying to attract non-native clients and are not willing to make any adjustment to accommodate them as the scoring criteria will only judge international test-takers how well they can adopt the English nativeness. In a simple sense, theoretical awareness is extant, but the practical application is missing. As Hamp-Lyons & Zhang (2001) stated two decades ago, “the question of which English(es) should be privileged on tests is particularly problematic and interesting in academic contexts where traditionally ‘standard’ forms of English are the only ones accepted.” (as cited in Davies et al., 2013, p.
Unfortunately, the issue is likely to remain unsettled for more decades to come, but is it for sure that these international tests will face more challenges mostly likely to come from the fields of World Englishes and ELF.

References
Arik, B. T., & Arik, E. (2014). The role and status of English in Turkish higher education. *English Today*, 30(4), 5-10. https://doi.org/10.1017/S0266078414000339

Berg, B. L. (2001). *Qualitative Research Methods for the Social Sciences* (4th ed.). Boston: Allyn and Bacon.

Blaxter, L., Hughes, C., & Tight, M. (2010). *How to Research* (4th ed.). London: Open University Press.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40.

Canagarajah, A. S. (2007). Lingua franca English, multilingual communities and language acquisition. *The Modern Language Journal*, 91(Focus Issue), 923-939.

Centre for Global Englishes Webpage (2018). About us | Centre for Global Englishes | University of Southampton. Southampton.ac.uk. Retrieved from http://www.southampton.ac.uk/cge/about/index.page

Davies, A., Hamp-Lyons, L. I. Z., & Kemp, C. (2003). Whose norms? International proficiency tests in English. *World Englishes*, 22(4), 571-584.

Dewey, M. (2015). Time to wake up some dogs! Shifting the culture of Language in ELT. In Y. Bayyurt, & S. Akcan (Eds.), *Current perspectives on pedagogy for English as a Lingua Franca. Developments in English as a Lingua Franca Franca [DELF]* (5) (pp. 121-134). De Gruyter Mouton, Berlin.

Galloway, N. (2013). Global Englishes and English language teaching (ELT) - Bridging the gap between theory and practice in a Japanese context. *System*, 41(3), 786-803. http://doi.org/10.1016/j.system.2013.07.019

Galloway, N., & Czarnecki, M. (2014). I get paid for my American accent: The journey of one multilingual English teacher (MET) in Japan. *Englishes in Practice*, 1(1), 1-30. http://doi.org/10.2478/eip-2014-0001

Galloway, N., & Rose, H. (2015). *Introducing Global Englishes*. Abingdon: Routledge.

Hamp-Lyons, L., & Zhang, W. B. (2001). World Englishes: Issues in and from academic writing assessment. In J. Flowerdew, & M. Peacock (Eds.). *Research Perspectives on English for Academic Purposes* (pp.101-116). Cambridge: Cambridge University Press,
Harding, L. (2012). *Language Testing, World Englishes and English as a Lingua Franca: The Case for Evidence-Based Change*. Invited key note address Centre for Internationalisation and Parallel Language Use Symposium 2012. Denmark: University of Copenhagen.

Hu, G. (2012). Assessing English as an international language. In L. Alsagoff, S. L. McKay, G. Hu, & W. A. Renandya (Eds.), *Principles and Practices for Teaching English as an International Language* (pp. 123-143). New York: Routledge.

Jenkins, J. (2014) *English as a Lingua Franca in the International University*. Oxford, New York: Routledge.

Jenkins, J. (2006). Points of view and blind spots: ELF and SLA. *International Journal of Applied Linguistics, 16*(2), 137-162. http://doi.org/10.1111/j.1473-4192.2006.00111.x

Jenkins, J. (2007). *English as a Lingua Franca: Attitude and Identity*. Oxford: Oxford University Press.

Jenkins, J. (2009). English as a lingua franca: Interpretations and attitudes. *World Englishes, 28*(2), 200-207.

Jenkins, J. (2011). Accommodating (to) ELF in the international university. *Journal of Pragmatics, 43*(4), 926-936. http://doi.org/10.1016/j.pragma.2010.05.011

Jenkins, J., Cogo, A., & Dewey, M. (2011). Review of developments in research into English as a lingua franca. *Language Teaching, 44*(03), 281-315. http://doi.org/10.1017/S0261444811000115

Jindapitak, N. & Teo, A. (2012). Thai tertiary English majors’ attitudes towards and awareness of World Englishes. *Journal of English Studies, 7*, 74-116.

Johnson, S., Milani, T. M., & Upton, C. (2010). Language ideological debates on the BBC “Voices” website: Hypermodality in theory and practice. In S. Johnson, & T. M. Milani (Eds.). *Language Ideologies and Media Discourse: Texts, Practices, Politics* (pp. 223-251). London: Continuum.

Jørgensen, M. W., & Phillips, L. (2002). *Discourse Analysis as Theory and Method*. London: Sage.

Kachru, B. (1992). *The Other Tongue: English Across Cultures* (2nd ed.). Urbana: University of Illinois Press.

Kachru, B. B. (1986). *The Alchemy of English: The Spread, Functions and Models of Non-Native Englishes*. London: Pergamon.

Khan, S. (2009). Imperialism of international tests: An EIL perspective. In F. Sharifian (Ed.), *English as an International Language: Perspectives and Pedagogical Issues* (pp. 190- 205). Multilingual Matters.

Knox, J. (2007) Visual-verbal communication on online newspaper home pages. *Visual Communication, 6*(1), 19-53.
Leung, C., Lewkowicz, J., & Jenkins, J. (2016). English for Academic Purposes: A need for remodelling. *Englishes in Practice, 3*(3), 55-73.

Lowenberg, P. (1993). Issues of validity in tests of English as a world language: Whose standards? *World Englishes, 12*(1), 95-106.

Luka, I., & Seniut, I. (2019). Developing students’ language competence and essential 21st century skills for future employability: The case of Latvia and Lithuania. *Acta Educationis Generalis, 9*(2), 1-23. https://doi.org/10.2478/atd-2019-0006

Matsuda, A. (2012). *Principles and Practices of Teaching English as an International Language*. Bristol, UK: Multilingual Matters.

McKay, S. L. (2002). *Teaching English as an International Language: Rethinking Goals and Approaches*. Oxford University Press.

McNamara, T. (2012). English as a lingua franca: The challenge for language testing. *Journal of English as a Lingua Franca, 1*(1), 199-202. https://doi.org/10.1515/jelf-2012-0013

Methitham, P. (2009). *An Exploration of Culturally-Based Assumptions Guiding ELT Practice in Thailand, a Non-Colonized Nation* (unpublished dissertation). Indiana University of Pennsylvania, Indiana.

Pauwels, L. (2012). A multimodal framework for analyzing websites as cultural expressions. *Journal of Computer-Mediated Communication 17*(3), 247-265.

Schreier, M. (2012). *Qualitative Content Analysis in Practice*. London: Sage.

Widdowson, H. (2012). ELF and the inconvenience of established concepts. *Journal of English as a Lingua Franca, 2*(1), 187-193.

Widdowson, H. G. (2015). ELF and the pragmatics of language variation. *Journal of English as a Lingua Franca, 4*, 359-372.