

Accent Difference Makes No Difference to Phoneme Acquisition

November 2022 – Volume 26, Number 3

<https://doi.org/10.55593/ej.26103a3>

Marc Jones

Toyo University, Japan
<jones056@toyo.jp>

Carolyn Blume

Technical University Dortmund, Germany
<carolyn.blume@tu-dortmund.de>

Abstract

ELT materials tend to use prestige variety speakers as models, an underlying assumption being that this is needed in order to acquire the phonology necessary to parse English speech (Rose & Galloway, 2019). Global Englishes Language Teaching (GELT) (Galloway & Rose, 2018) provides the potential for movement away from such ‘native speaker’ ideologies, but lacks empirical evidence. In this study, the use of GELT input in comparison with prestige varieties of English was investigated. Sixteen first-year L1 Japanese university students in an English Medium Instruction programme participated in a self-paced listening study via a learning management system (LMS). All participants were tested on their perception of the English vowels /æ/, /ʌ/, /ɜ:/ and /ɔ:/. After this pretest, they were separated into two groups: using edited TED talks, the experimental group (G) (N=8) watched videos of Global English varieties, and the control group (P) (N=8) watched videos of prestige English varieties. Both groups acquired losses, i.e., immediate posttest scores were mainly lower than pretest scores on vowel identification. Scores were predicted by the variation in interval between lessons and posttest, but not by the varieties of English used. This provides support for the view that GELT is as valid a language teaching approach as using prestige varieties.

Keywords: Global Englishes, listening, multimodality, phonology

Global Englishes and Global Englishes Language Teaching (GELT) (Galloway & Rose, 2017) are an increasingly important part of the English Language Teaching field, gaining prominence as the problematic nature of concepts such as ‘native speakerism’ (Aneja, 2016; Holliday, 2006; Kiczkowiak, 2019; Lowe & Kiczkowiak, 2019) become subject to growing awareness. Moreover, given the increasing availability of authentic multimodal resources that present a range of English varieties, and learners’ exposure to them, GELT potentially promises both a philosophical and pragmatic transformation in English courses of study. However, there is little empirical data concerning how the use of Global Englishes affects language learners’ acquisition in measurable terms. Without a better understanding of what linguistic challenges and potential benefits may be associated with such approaches, the development of appropriate pedagogy is not possible.

In this paper, we examine how Global Englishes compare to the use of prestige varieties of English in a self-paced multimodal listening course at a Japanese university in an English Medium Instruction (EMI) course. Stemming from philosophical stances regarding Global Englishes, linguistic models of phonological acquisition, and instructional approaches to multimodal language learning, this contribution provides a rationale for the pedagogical approach employed here while offering empirical analysis of its implementation. While the results do not demonstrate an empirical advantage towards using multimodal Global English materials regarding phonological acquisition, they likewise indicate that there is no marked disadvantage to doing so. In light of the substantial philosophical and conceptual benefits associated with using Global Englishes, we thus argue that these data make a strong case in favour of its utilization.

This paper begins by describing relevant arguments in favour of GELT and extant research that examines attitudes towards prestige and non-prestige varieties of English among English language learners. Highlighting the research gaps in this area regarding acquisition, we will consider empirical data regarding phonological acquisition of prestige and non-prestige varieties. The potential of TED talks as multimodal language learning resources with elements that pertain to Global Englishes will be summarized, based on the literature, before describing the ecological study that was implemented here. Following an explanation of the structure of the learning opportunity and the study design, we will use Bayesian analysis to adequately address the findings in relation to the given sample size. These results will be contextualized in a subsequent discussion before we conclude with an outlook for further areas of inquiry.

Global English Language Teaching: Theory and acquisition

Global English Language Teaching: The theoretical argument(s)

In research literature, definitions of L2 varieties are given in comparison to so-called ‘standard varieties,’ hereafter referred to as ‘standardised varieties’ to emphasize our understanding that their use as benchmark varieties is socially constructed. While standardised varieties are generally accorded prestige, other varieties – especially English ones – may be regarded as deficient (Kubota, 2015). This accordance of prestige is linked to the hegemonic status of various nation states, the economic wealth of those states, and also the social capital afforded to citizens of those states or those who may pass as such citizens. Within nation states, further distinctions of prestige accorded to various varieties may also be found. In other words, the prestige attached to language varieties is socially constructed, and depends upon the community in which a given variety is used. This explains in large part why “ranked uppermost among these varieties in terms of prestige and privilege are the two metropolitan Englishes based largely in London and US cities like New York and Washington” (Rubdy, 2015, p. 45). Likewise, extant research examining learners’ attitudes towards these so-called non-

standardized Englishes report biases among learners against them, especially as regards accents (e.g. Jacobsen, 2015; Meer, Hartmann & Rumlich, 2021).

In this article we use the term Global Englishes in line with its adoption by Rose & Galloway (2019), who use it as an “inclusive term” (p. 3) to “consolidate research in World Englishes, English as a lingua franca and English as an international language” (p. 4). In short, Global Englishes is a term that refers to use of the language without reference to the problematic notion of nativeness (Holliday, 2006). Given the fact that Global Englishes are largely absent from most commercial ELT materials (Kiczkowiak, 2021), they are also probably underrepresented in practice, leading to a lack of familiarity with such forms that consequently perpetuates discrimination against the speakers of these varieties. This sets up a vicious cycle, ultimately resulting in learners who are ill-prepared for encounters with speakers of non-prestige varieties of English, who view these speakers through a deficit lens (Kubota 2015), and who – depending on their own positionality – likely experience feelings of inefficacy regarding their own competence in their additional language. This has potentially significant ramifications that extend beyond the language learning classroom. As Piller (2016) remarked, “perceiving peripheral speakers as incomprehensible licenses their exclusion from participation in development contexts” (p. 199). In other words, labelling Global Englishes speakers as incomprehensible deskills these speakers and learners, and deprofessionalizes these educators. In English medium instruction (EMI) settings, this attitude sustains a neocolonialist imperative for teachers who are prestige variety speakers and correspondingly marginalises non-prestige speakers of Global Englishes.

Global Englishes accents and phonological acquisition

While advocates of GELT have, for all the aforementioned reasons, been vocal about its theoretical necessity as a corrective to standardised hegemonic language forms, pragmatic issues concerning acquisition have garnered less attention. One of the areas in which varieties are identified is in the phonetic manifestations of their phonology, namely accent. Derwing and Munro (2009) defined speakers’ accents as “the ways in which their speech differs from that local variety of English *and* the impact of that difference on speakers and listeners” (p. 479; italics added). In this comparative approach, it is not necessarily the case that a ‘foreign accent’ is defined in terms of deviation from a prestige accent, but rather in terms of *deviation from what a listener is accustomed to*, or the ease with which a listener can comprehend speech. It is obvious that familiarity contributes to intelligibility. Carey, Mannell and Dunn (2011), for example, investigated ratings by speaking examiners in standardised tests and found that “pronunciation was rated higher by a significant proportion of OPI [oral proficiency interview] examiners when the candidate’s interlanguage phonology was familiar and lower when it was unfamiliar” (p. 215, added parentheses). It is therefore clear that exposure to a number of speakers to increase intelligibility is a necessary, but not sufficient, element informing comprehension (Varonis & Gass, 1982). In addition to the linguistic features of grammar and pronunciation, social variables come into play. This includes the racialization of speakers, with a resulting hierarchy of mainly ‘white’, ‘Western’, ‘native speakers’ and an accompanying marginalisation of racialized speakers as ‘non-native’ (Aneja, 2016). It becomes clear that both speech and image, as well as familiarity, interact to inform perceived intelligibility of oral Global Englishes.

By categorising perceived acoustic phenomena as phonemes, listeners can potentially parse the speech of interlocutors or speech in media. Various models of how this occurs among L2 learners have been developed and tested over decades, with the PAM (L2) model developed by Best and Tyler (2007) most relevant to L2 learners of English living outside L1 English speech communities. The PAM (L2) posits that learners will categorise L2 phones according

to L1 phonemes, and then at a later stage, change this categorisation through creation of categories for L2 phonemes. The speed at which this occurs is dependent upon individual differences, but also relies on sufficient L2 input, with ‘sufficient’ being defined as a length of time beyond 6-18 months in a L2 environment (Best & Tyler, 2007, p. 21). While research is sparse, it seems that such parameters might hold true for learners to recognize non-standardised varieties as familiar ones.

Multimodal Global English resources and phonological acquisition

Use of available, multimodal Global English resources offers a potential, partial solution to the absence of Global Englishes in commercial materials that contributes to the general lack of familiarity English language learners report regarding these varieties. The popularity of online video platforms such as Youtube and Netflix, as well as TED Talks in particular, as discussed in more detail below, has led to easy access to media in many languages, and also Englishes from many regions. By using easily accessible Global Englishes for language input, the unfamiliarity of varieties underrepresented in ELT can be addressed. Whether this pedagogical approach is sound in terms of phonological acquisition, however, remains largely unexplored. One exception is that Ockey and French (2016) found that self-reported familiarity with British accents correlated with higher comprehension, but did not observe the same for familiarity with Australian accents. Additionally, Saito et al. (2019) found that closeness of listeners’ L1 to a L2 speaker’s L1 correlated with increased comprehensibility, in keeping with previous findings by Bent and Bradlow (2003). However, vowel acquisition by Japanese L1 users was not investigated in the above studies. Furthermore, the L1 of speakers of the English varieties to be used are all distant from that of our (Japanese) population.

Multimodal phonological acquisition

The language that we use and intend for learners to acquire is multimodal; the auditory information is heard while seeing articulatory gestures (mouth movements) and paralinguistic gestures (other body movements such as pointing). This may aid in phonology acquisition and language learning processes more generally. Ravelli (2018) asserted that “there is a continuum from these inherent forms of multimodality – that is, the visual presence of written language, and the physical presence of spoken language – to the related phenomenon of explicit co-configuration of multiple modes” (p. 434). That is, language events are not subject to a binary of multimodality or single modality but should be considered according to what extent they consist of different modalities. For example, a TED video consists of the speaker’s speech, their facial expression and articulatory gestures, whole body movements, audience reaction shots, slide presentations with on-screen text and other modes of communication and therefore may be considered as highly multimodal. When people without sensory impairments or processing issues encounter interlocutors, their attention is focused primarily on the interlocutor’s face. This means that visual and linguistic salience “assign a property to a visual/linguistic unit that renders it perceptually more prominent in an array of competing units, which is crucial in cases where selective attention is useful or necessary” (Rącz, 2013, p. 23). Suvorov (2015), for example, found that “L2 test takers spent statistically significantly more time watching content videos (58% of the total time the videos were shown) than context videos (51%)” (p. 476), where content videos are those showing the speaker and their gestures, whereas context videos were those that presented visual information as text or graphics along with speech. This could mean that the participants were focused on the people, including the faces and speech articulation, more often than other visual contexts. While it could also be the case that the content videos in Suvorov’s study were simply more interesting than the context videos – a claim that Suvorov himself makes (p.477) – the findings echo those of other interventions that highlight the role of visual reinforcements of articulation processes. Batty (2020) similarly

found, using eye-tracking measures and participant interviews, that learners spent more gazing time on faces than on bodily gestures in video-mediated listening.

In computer-assisted language learning (CALL), multimodality is theorized to be beneficial based on assumptions regarding dual-coding. Mayer (2021) argues that simultaneous processing of visual and auditory information leads to more efficacious encoding. In addition to visual articulatory speech information, paralinguistic features such as gestures, distance and/or physical contact between interlocutors, and prosody further aid comprehension (Hoven, 1999). However, orthographic information may actually interfere with phonological acquisition where L1 and L2 phonology and orthography pairings are incongruent (Hayes Harb & Barrios, 2021; Barrios & Hayes Harb, 2020; Mathieu, 2016), which means closed captions may not be beneficial for this purpose. Therefore, when using multimodal input one must ensure that the different modalities work to facilitate the second language acquisition process rather than interfere. In particular, if there is a potential interlingual interference, caption use should generally be avoided.

Although early uses of CALL for perceptual training (training the ability to perceive differences in speech sounds) reflected theories of learning predicated on repetitive exposure to ‘native-like’ pronunciation, subsequent technological and conceptual developments have led to an interest in models where learners have greater levels of interaction with technology, as well as greater diversity of speakers in media used (Blake, 2016; Davies et al., 2013; Hubbard, 2017). However, in addition to becoming arguably more inclusive of non-prestige speech varieties, these approaches have tended to focus more on global listening strategies (Hoven, 1999; Jolley & Perez, 2020). Coupled with a potential lack of teacher knowledge about, and skills for, addressing listening skills at the phoneme level (Jones, 2020), research and pedagogy that integrate authentic (here operationalised as materials produced by English speakers, primarily for purposes other than language learning; see Blume, 2022, for more detailed analysis) audiovisual texts for segmental listening skills over a range of speech varieties is sparse.

Multimodal TED talks for listening comprehension of Global Englishes

As in this study, TED (Technology, Entertainment and Design) talks have previously been used in ELT settings with learners of varying L1 backgrounds and level of English knowledge as one source of audiovisual input (Coxhead & Bytheway, 2014; Madarbakus-Ring, 2020; Montero Perez, 2019). TED talks are attractive resources due to the range of topics they address, their easy accessibility, and their effective use of principles of public speaking. They are also clearly multimodal artefacts, consisting of web page, video, options within the video playback interface, spoken text, speaker’s gesture and camera angle, as well as the interactive transcript, which allows skipping to the relevant section of the recording when clicked, and/or subtitles. In small scale studies, learners’ affective reception of TED talks has been generally high, due to their perceived authenticity (Ahluwalia, 2018; Takaesu, 2017, p. 114) and the range of subjects from which listeners can choose based on their own intrinsic interests (Kozłowska, 2021, p. 210; Wu, 2020, p. 33). Additionally, due to the camera angles and production values of TED talks, “Global Englishes speakers are invariably displayed favorably, namely as an expert to be taken seriously regardless of national and/or linguistic backgrounds” (Schildhauer et al., 2021, p. 203). By setting Global English speakers as being of the same status as prestige variety speakers, TED talks can help to legitimize the presence of GELT.

While the library of over 5000 TED talks (<https://www.ted.com/>) includes speakers of many English varieties, existing descriptions of the talks used in the aforementioned studies do not indicate the varieties with which the studies were carried out. Data collected in these studies reveal that the English learners regularly report that they found the TED talks difficult to

comprehend, at least initially. However, without further descriptions of the specific TED talks utilized by the learners, few conclusions can be drawn about the effect of various varieties on comprehensibility. The reported data suggests that learners found both prestige varieties and varieties from less prestigious and less familiar contexts challenging (Takaesu, 2017; Kozińska, 2021). At the same time, the learners expressed an interest in listening to TED talks with non-prestige varieties, finding both the range of topics and diversity of the speakers motivating (Takaesu, 2017, p. 114; Kozińska, 2021, p. 210).

Although existing studies thus suggest that TED talks may be efficacious for language learning, due to both their multimodality and their affective impact, this prior research is limited in its scope to self-reports of learners' perceived listening facility and attitudinal factors. Nor does it identify the English language varieties that learners are exposed to. Research that empirically examines TED talks, as a source of multimodal input for non-prestige varieties on measures of phonological acquisition, is absent. While a great deal of research has been done in the areas of phonological acquisition and multimodal language learning resources, few ecological studies use authentic materials to evaluate and address learners' phonological acquisition. This gap between theory and practice exists despite the fact that implementing established models of phonology acquisition can be pragmatically realized in multimodal learning opportunities, regardless of the language varieties being targeted.

Methodology

In this paper, we describe a study that examines how Global Englishes compare to the use of prestige varieties of English in a self-paced multimodal listening course. The self-paced course was embedded in an EMI course at a Japanese university. The participants were drawn from the largely L1 Japanese sector of the cohort (approximately 70 percent), with international students making up the remaining 30 percent of each cohort. A convenience sample was taken from the first author's class, based upon informed consent and completion of the self-paced listening course described below. We aim to address the following research questions:

RQ1: How does GELT affect receptive phonology acquisition of vowels?

RQ2: How does the interval between learning sessions contribute to phonology learning gains?

While the sample size is small, the ecological validity of the intervention is one measure of its validity, with statistical methods selected to accommodate the size of the population. Instead of null hypothesis significance testing using frequentist statistics, Bayes factor analyses were run as more rigorous applications for smaller sample sizes (see Norouziyan, de Miranda & Plonsky, 2019). Additionally, generalized linear models (Baayen, Davidson & Bates, 2008, p. 391) with Monte Carlo Markov chain (MCMC) sampling (Thomopoulos, 2013) were used to mitigate power problems typically associated with small samples (for a detailed view see van de Schoot & Miočević, 2020). After analysis, we provide a discussion of the results found and then provide a conclusion along with a comment on the limitations of the study.

Procedures

This study was undertaken with an intact class taught by the first author at a private university in Japan. No learners reported hearing problems or disabilities that affect speech reception or processing. The learners are of approximately equal proficiency level, in that the goal for the end of the first year of their programme of study is to attain an IELTS score of 6.0 or above. This is roughly intermediate level, or deemed sufficient by higher education institutions for entry to an exchange programme. Learners took a pretest and then were separated into two groups paired roughly by score, i.e. two learners gaining the same or almost the same score

were assigned to different groups to create balance across conditions. From the intact class, sixteen learners returned consent forms permitting their data to be used for this study, therefore the sample size was N=16, of which eight students were in each condition described below. One learner reported Chinese L1 but Japanese as a dominant language, and was placed in the experimental condition group described below. The learners gained academic credit for completion and partial completion of the self-paced listening course described below. No financial rewards were offered or given.

Tests

Learners listened to 32 CVC words recorded by the first author using a Blue Snowball iCE microphone and PRAAT software (Boersma & Weenink, 1992), eight each of the vowels /æ/, /ʌ/, /ɔ:/, and /ɜ:/, labelled in plots and tables as TRAP, STRUT, THOUGH and NURSE respectively. (See the full word list in Appendix 3). The reason for focusing on vowels is that Japanese does not have as many central vowels as English (Keating & Huffman, 1984) and therefore during listening, the inability to accurately perceive or discriminate these English vowels may cause comprehension difficulties. Additionally, the selected vowels have similar formant ('undertone') values: /æ/, /ʌ/, and /ɜ:/ have similar first and second formant values, and all vowels have similar third formant values according to Deterding (1997, p. 49), meaning that while classification for L1 users may be straightforward, the differences between them may not be salient enough for easy L2 acquisition. Furthermore, given the breadth of work on consonant acquisition (for example Sheldon & Strange, 1982; Lambacher, 1999; Lotto, Sato & Diehl, 2004; Sueyoshi & Hardison, 2005) in comparison to the scant work on vowel acquisition for L1 Japanese learners of L2+ English, (Ingram & Park, 1997; Nishi & Kewley-Port 2005, 2008), more work is warranted.

The test was the same for both pretest and posttest conditions, with different question sequences in the pretest and posttests. The didactic materials and the tests were delivered through a learning management system (Moodle, 2020), and the vowels identified by choosing a picture of a common noun that shares the same vowel as the recorded word. The pictures provided were of the words *cat*, *sun*, *door*, and *bird* for the vowels /æ/, /ʌ/, /ɔ:/, and /ɜ:/ respectively.

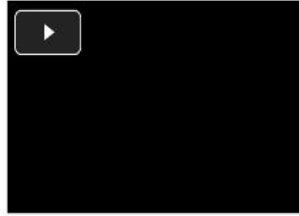
Self-paced listening course

The course content was provided through the lesson module in Moodle. Each lesson consisted of four sets of a vowel identification exercise in the context of a short word (CVC, CVCC or CCVC) with automated feedback (Figure 1), then an utterance decoding task containing the word from the vowel identification (Figure 2). This provided not only vowel discrimination as a discrete activity but integrated it into a realistic decoding task, providing affordances for participants to integrate discrimination into their decoding and parsing processes. Participants then watched a 10-minute excerpt of the TED talk and subsequently provided a summary. Captions were not available due to the assumption of phonological interference given the incongruence between English orthographic and Japanese L1 phonological systems as discussed previously. This structure was repeated over five lessons, and was intended to be completed on a weekly basis, for a total of five weeks. Group P listened to videos of speakers of standardised prestige English varieties, while Group G listened to videos of speakers of Global Englishes, detailed in Table 1, below.

Lesson 1 G

You will watch some parts of a video, identify words, and write short summaries, as well as give your reaction.

Choose the picture of the word with the vowel that matches what the speaker says.



Submit

Figure 1 Screenshot of vowel identification task

Lesson 1 G

You will watch some parts of a video, identify words, and write short summaries, as well as give your reaction.

Write what the speaker says.



Your answer

Submit

Figure 2 Screenshot of utterance decoding task

Table 1. Table of TED talks used for each lesson. Group G - Global Englishes; Group N – Prestige ‘Native’ Englishes. Variety of English or additional language influence on English given after year. Varieties only defined by L1 influence and not dialect or specific region.

Theme	Group G Talk	Group N Talk
Urbanization	Adegbeye (2017) Nigerian	Speck (2013) American
Maternal healthcare	Hegde (2021) Indian	Howell (2018) American
Environmentalism	Jun (2021) Chinese	Francis (2019) English
Personal finance	Belle (2021) Kenyan	Gibson (2019) American
Political disagreement	Cekic (2018) Turkish/Danish	Pearlman (2019) American

Results

Table 2. Mean pretest and posttest scores, time (in days) from L1-L2, L3-L4, L4-L5, L5-post, and mean standard deviation of intervals.

Condition	Mean Pre	Mean Post	Mean Interval L1_2 (Days)	Mean Interval L2_3 (Days)	Mean Interval L3_4 (Days)	Mean Interval L4_5 (Days)	Mean Interval L5_Post (Days)	Mean Interval SD
G	26.5	24.50	16.37	2.31	2.72	6.86	4.83	688.62
N	27.5	24.37	4.38	6.49	1.93	4.99	9.29	170.50

Pretest and posttest scores were used to calculate gains for each participant in R (R Core Team, 2020), both overall and for each vowel studied. These were then analysed in a Bayesian Welch's t-test using BEST (Kruschke & Meredith, 2021) for two-sample comparison, i.e., cross-group comparison. Beyond this, GLMs were constructed in Stan (Guo et al., 2020) and Rstanarm (Gabry & Goodrich, 2020) with MCMC sampling in order to find what factors contributed to the resultant scores. These models were then compared using the bridgesampling (Gronau et al., 2020) and BEST (Kruschke & Meredith, 2021) packages. Bayesian priors were set at null (i.e. default broad priors) due to the lack of similar studies and available prior evidence to provide a suitable prediction.

As can be seen in Table 2, there were negative gains for both groups P and G on average, with a minority of learners making minimal positive gains. Gains in group G were generally seen through greater acquisition of TRAP while gains in group P were made in acquisition of NURSE (Table 3). However, using the guidelines for interpreting Bayes factors in Norouzian, de Miranda and Plonsky (2019, p. 252) both these gains are marginal and may be attributed to chance. This is reinforced by the Bayes factors for all the gains.

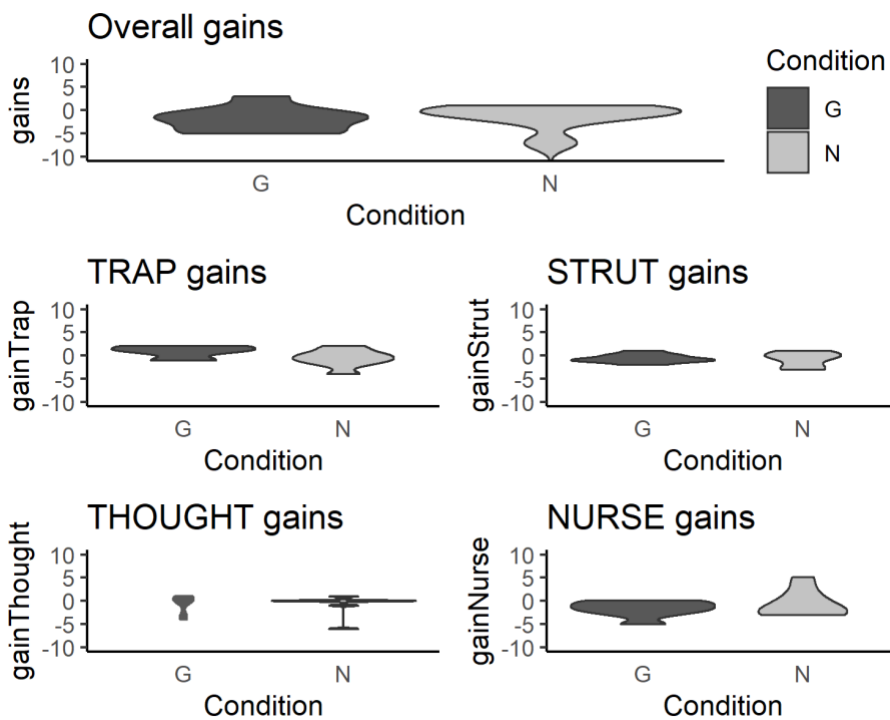


Figure 3. Mean gains by group overall and by vowel

Table 3. t-tests for group G vs. group P

Description	t Score	Bayes Factor
Overall gains	0.50	0.32
TRAP gains	1.90	0.22
STRUT gains	0.39	0.77
THOUGHT gains	0.24	0.77
NURSE gains	-0.95	3.08

When analysing the GLMs to account for the trend in gains, it appears that condition is irrelevant, as seen in Table 4. The increase in Bayesian factor between ‘Gains, condition, participant and interval standard deviation: no interaction vs interaction’ and ‘Gains, participant and interval standard deviation: no interaction vs interaction’ shows that the largest interaction is due to the standard deviation in time intervals between lessons and posttest. Due to analysis of linear models in comparison with one another, ceiling effects are unlikely to be present, as can be observed in the series of similar linear models being compared.

Table 4. Bayes factors for the comparisons of different GLMs

Description	Bayes Factor
Gains, condition and intervals no interaction vs interaction	Inf
Gains, condition, participant and interval standard deviation: no interaction vs interaction	0.45
Gains, participant and interval standard deviation: no interaction vs interaction	118.91
Gains, condition, participant and individual vowel gains: no interaction vs interaction	0.00
TRAP gains, condition, participant and interval standard deviation: no interaction vs interaction	0.002
STRUT gains, condition, participant and interval standard deviation: no interaction vs interaction	0.64
THOUGHT gains, condition, participant and interval standard deviation: no interaction vs interaction	1.10
NURSE gains, condition, participant and interval standard deviation: no interaction vs interaction	0.18

Discussion

As can be seen from the results of both the MCMC sampled GLM and the Bayesian t-test, according to Nourozian et al.’s (2019) guidelines for Bayes factor interpretations (p. 252), results of the t-tests show that there is insufficient evidence to support the use of prestige varieties of English over Global English, with relatively low t scores and Bayes factors very close to suggesting only anecdotal evidence strength. Therefore, we can state that there is effectively no difference in phonology acquisition between groups that listened to prestige variety Englishes and non-prestige Global Englishes. The only factor that seemed to affect phonology acquisition was the standard deviation of intervals between lessons in the self-paced course. The evidence thus suggests that GELT is no less effective an approach than teaching using prestige models in a self-paced listening course based on TED talks. The learners using Global English materials did not have significant differences or worse losses than the prestige variety group. While this may seem intuitively or experientially obvious to practitioners using GELT in their day-to-day teaching, there is little so far in the extant literature to provide empirical evidence for the benefits of GELT. This study, despite its relatively small scale, thus offers data to support a GELT approach, in that it is not inferior to the use of prestige models.

In addition to marginalizing both teachers and learners who do not reflect stereotypes of so-called ‘native speakers,’ a prescriptivist, prestige-based standard holds as an ideal a largely unattainable goal that is at odds with contemporary notions of communicative competence. Rather, given patterns in global deterritorialisation (see Jacquemet, 2005, p. 261), the ability to communicate transculturally is of utmost importance. Of course, whether using prestige or non-prestige varieties, transcultural communication is possible. Our point here is that, due to prestige varieties having already been widely used in ELT, parsing of prestige varieties poses fewer problems for learners whereas less prestigious Global Englishes have been absent from materials and this lack of familiarity may cause problems. Ensuring that learners are equipped to communicate among an increasingly mobile global population means that an understanding of English as it is used as a ‘lingua franca’ (ELF) (Jenkins, 2000; Seidlhofer, 2011) thus needs to be both a philosophical and a didactic priority. (In this context, we use the term Global Englishes to refer to what speakers say, whereas ELF refers to the use of English as a common language among speakers of different languages.) While the dominance of English in transnational academic, popular, and economic contexts admittedly perpetuates a problematic neoliberal agenda (Jacobsen, 2015; Kubota, 2015, 2021), fostering learners’ ability to function in this context offers them a degree of agency. That is, by teaching not only with prestige varieties but also non-prestige varieties, learners gain skills to communicate more effectively across a wider array of communities using English.

In light of the fact that the absence of regular intervals is associated with the lack of phonological acquisition, more research is needed to examine the role played by students’ proclivity to complete all the activities at once, as revealed by completion data obtained via the learning management system. It is plausible that progressing in irregular increments throughout the course of study might inform the lack of gains, regardless of Global English variety. In keeping with research regarding practice for L2 vocabulary retention (Bloom & Shuell, 1981) and L2 syntax (Bird, 2010), regular attention over a period of time might be more likely to contribute to measurable gains, though this is yet to be confirmed for phonology. If this is the case, it would be important to address the pedagogical issues involved in managing student progress in a self-paced course in particular contexts.

Alternatively, more intensive or more prolonged learning opportunities might be necessary to facilitate gains in vowel perception, although this is currently unclear. As suggested by studies about familiarity with accents more generally (cf. Varonis & Gass 1982) and discussed above, it is unclear what amount of exposure is necessary to improve phonological acquisition. Due to the easier measurement of pronunciation, this appears to frequently be a proxy measure of perceptual phonology. Tyler & Best (2007) suggested that research populations of language learners tend not to differ greatly in their phonology after 18 months living in the community where they study, although their focus, to be clear, is on pronunciation. Mora and Fullana (2007) found that length of study does not affect Catalan/Spanish bilinguals’ perception of English /æ/-/ʌ/ or /i/-/ɪ/ contrasts (pp. 1615-1616). As summed up by Flege and Bohn (2020), “[i]t is unknown at present how much L2 input is needed to form phonetic categories in an L2 and optimally adapt them to everyday use. This may depend, at least in part, on the uniformity of the L2 speech input that is received” (p.13). In short, the time required for L2 vowel acquisition may be subject to individual differences, such as time spent within different languages, but existing research is scant and contradictory. More research in this area would be useful to all speech learning researchers, whether related to English or other languages.

The question also remains whether these particular videos, or videos in general, are unsuitable media for phonological acquisition. McCrocklin (2012) demonstrated that there was no significant effect for video over audio for /i/ - /ɪ/ contrasts in an ESL class in the USA with a large number of L1 Chinese speakers. However, Navarra and Soto-Faraco (2007) found that

video was more beneficial for perceiving /e/-/ɛ/ contrasts than audio alone. Furthermore, Becker and Sturm (2015), while not reaching statistical significance in comparing audio and visual modes for L2 French listening comprehension, found that video had a larger effect size than audio only. Other factors that might hypothetically affect the suitability of the media could, instead, lie in the complexity or unfamiliarity of the topics or length of the video excerpts, which ran to approximately 10 minutes, with very short excerpts preceding the final full 10-minute viewing. Unfamiliarity with listening tasks without the affordance of negotiating meaning with the speaker may also be possible factors contributing to our results.

We believed that using videos for listening comprehension would facilitate acquisition through greater visual salience. One reason for this may be that language processing seems to be linked to the articulatory features. That is, when visual articulatory information is perceived, it is processed in the same area of the brain as auditory information, Broca's region (Glanz et al., 2018). This means that if auditory input can be made more visually salient, or that learners have more than one source of evidence for the phonetic information they are attempting to parse, they are more likely to be successful in their efforts. This appears to be the case in studies by Hardison (2003; Sueyoshi and Hardison, 2005), where consonants were acquired more readily by learners exposed to audiovisual information than those exposed to audio-only information. Inceoglu (2022), on the other hand, found mixed results of an audiovisual effect in the perception of L2 French vowels, outcomes being mediated by prior word knowledge and individual vowel features. This is potentially due to the nasality of some French vowels in contrast to the absence of nasal vowels in English and, which may contribute to a lack of visual salience among the former. Like Inceoglu, our results suggest that the videos do not significantly contribute to acquisition. As we did not have an audio-only condition, direct comparisons are difficult. However, it may be that audiovisual effects vary for consonants and vowels, regardless of language.

It appears likely that viewing multiple videos shortly before the end of their course of study could be one reason for the losses in perceptual discrimination of the target vowels. This could be attributed to working memory fatigue or waning selective attention. It may also be the case that this loss is part of a developmental trajectory, similar to the 'U-shaped' acquisition trajectory theorised by Shirai in relation to lexical learning (1990). Whether the processes and theories postulated regarding 'U-shaped' trajectories for initial language acquisition and second language semantic items are relevant for phonological acquisition has not been established.

Limitations

One of the clearest limitations of this study is the small sample size, and while this is somewhat mitigated by the use of Bayesian statistical approaches and MCMC sampling of the data in creating the GLM, the effect sizes reported should be approached with caution. While ecologically valid, the study is only one intervention with a relatively small scope. More work is required to account for the lack of phonology gains among participants. While other factors might play a role, it seems that the self-paced nature of the course led to some students 'cramming' the practice into unevenly spaced, and brief intervals, to the detriment of their phonology acquisition. This behaviour might reflect the timing of the sequence at the end of the academic year, fatigue from multiple competing obligations due to the close of the semester, or the unique situation still informed by COVID-19-related restrictions and extensive digital interactions. Such 'limitations' are, in the context of a study embedded in an intact language learning setting, not to be considered flaws in the research design. Rather, they highlight the human factors that play into formal language learning in a classroom setting that cannot ever be eliminated and should therefore not be excluded from empirical studies.

At the same time, ecological validity in one context does not necessarily result in generalisability across other contexts. Therefore, we welcome more research by scholars working in different contexts to build on the methodology in this study. This would develop a broader evidence base for phonology acquisition studies in pedagogical rather than laboratory-based work. Additionally, it would be highly useful for the field of ELT in general if further investigations into the use of GELT for second language acquisition were undertaken. GELT is becoming more important in ELT as the effects of globalisation and the exposure to less hegemonically dominant English varieties becomes a necessity rather than a curiosity limited to the university classroom.

Conclusion

While many questions are raised from the project, we would like to believe that we have contributed somewhat in showing that GELT is not a lesser approach to English teaching than the prestige-variety status quo. In short, teachers should not be reluctant to use a range of English varieties in their classroom practice.

While the outcomes of this study can be considered positive as regards GELT, they are sobering in light of the lack of gains made regarding receptive phonology among both groups. According to the data, the lack of gains can be attributed to the lack of regular intervals between undertaking lessons for many of the learners. While adult learners may be expected to be able to make appropriate decisions about their own learning, the patterns of interaction with the material in the LMS suggest that this is not universally the case, at least among this cohort. Given the effect on learning outcomes, as evidenced by this admittedly small study, more attention needs to be given to balancing the pedagogical need for effective monitoring with young adult learners' autonomy-related needs. The data showing how pacing interacts with acquisition clearly illustrate the ways in which second language acquisition research and didactic design need to go hand-in-hand.

Regardless of these caveats, it is clear that phonology acquisition is necessary for listening, because acoustic information needs to be parsed in order to be meaningful, and that this can demonstrably occur through input provided by speakers of any variety of English. Indeed, it seems apparent that a GELT approach (Galloway & Rose, 2014; Rose & Galloway, 2019) would be more appropriate than the status quo, not only for purposes of increasing language acquisition, but also for the purposes of reducing implicit colonialist, 'native speaker' ideologies in teaching. As Derwing and Munro (2014, p. 219) pointed out in reference to communication between L1 users and conversational participants with alternative L1s, the responsibility for intelligible communication resides with both interlocutors. At the same time, they made it clear that "[b]ecause the activation of biases is a phenomenon that takes place within the listener, addressing its effects requires changes in the listener rather than the speaker. Just as one cannot legitimately expect the target of a racist act to correct the problem by changing color, one cannot expect a non-native speaker to 'drop' an accent because others respond negatively to it. Rather, ample evidence indicates that accented speech is a normal aspect of second language development, and there is no evidence that accents can be routinely eliminated through pedagogical intervention" (2014, p. 221). As such, we disagree with points of view such as that of Tschirner (2011), who stated that, in pedagogical designs, "speakers or language that is difficult to understand" should be excluded (p. 35), since this begs the question of how these speakers or the language could ever become easily understood if they are, or it is, perpetually excluded. Seen in this light, addressing learners' perception of Global English pronunciations must combine efficacious phonological training with pedagogical measures to address potentially discriminatory attitudes.

About the Authors

Marc Jones is a language teacher and researcher at Toyo University. His research interests are primarily listening and phonology, extending to various other aspects of language teaching. ORCID ID: 0000-0002-2004-1809

Carolyn Blume holds the chair for digitally-mediated teaching and learning in the Dortmund Competence Center for Teacher Education & Research (DoKoLL) at the Technical University Dortmund, Germany. A member by courtesy of the Faculty of Cultural Studies (English), Dr. Blume's research focuses on initial and ongoing teacher education for English language teaching, specifically in the areas of inclusion and diversity, and digital media and digitality. Dr. Blume is the 2022 recipient of the IDEA Award, honoring university level teaching that is inclusive and that prepares teachers for diversity in their classrooms. ORCID ID: 0000-0002-4788-593X

Acknowledgements

Support for this article was provided in part by DoProfil, part of the "*Qualitätsoffensive Lehrerbildung*," a joint initiative of the Federal Government and the *Länder* which aims to improve the quality of teacher training. The programme is funded by the Federal Ministry of Education and Research. The authors are responsible for the content of this publication.

To Cite this Article

Jones, M. & Blume, C. (2022). Accent Difference Makes No Difference to Phoneme Acquisition. *Teaching English as a Second Language Electronic Journal (TESL-EJ)*, 26 (3). <https://doi.org/10.55593/ej.26103a3>

References

- Ahluwalia, G. (2018). Students' perceptions on the use of TED Talks for English language learning. *Language in India*, 18(12), 80-86.
- Aneja, G. A. (2016). Rethinking nativeness: Toward a dynamic paradigm of (non)native speaking. *Critical Inquiry in Language Studies*, 13(4), 351-379. <https://doi.org/10.1080/15427587.2016.1185373>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Barrios, S. L., & Hayes-Harb, R. (2020). Second language learning of phonological alternations with and without orthographic input: Evidence from the acquisition of a German-like voicing alternation. *Applied Psycholinguistics*, 41(3), 517-545. <https://doi.org/10.1017/S0142716420000077>
- Batty, A. O. (2020). An eye-tracking study of attention to visual cues in L2 listening tests. *Language Testing*, 026553222095150. <https://journals.sagepub.com/doi/10.1177/0265532220951504>
- Becker, S., & Sturm, J. L. (2017). Effects of audiovisual media on L2 listening comprehension: A preliminary study in French. *CALICO Journal*, 34(2), 147-177. <https://doi.org/10.1558/cj.26754>
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, 114(3), 1600-1610. <https://doi.org/10.1121/1.1603234>

- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language Experience in second language speech learning: In honor of James Emil Flege* (Vol. 17, pp. 13–34). John Benjamins Publishing Company.
<https://doi.org/10.1075/llt.17.07bes>
- Bird, S. (2010). Effects of distributed practice on the acquisition of second language English syntax. *Applied Psycholinguistics*, 31(4), 635–650.
<https://doi.org/10.1017/S0142716410000172>
- Blake, R. (2016). Technology and the four skills. *Language Learning & Technology*, 20(2), 129–142. <http://dx.doi.org/10125/44465>
- Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of second-language vocabulary. *Journal of Educational Research*, 74(4), 245–248. <https://doi.org/10.1080/00220671.1981.10885317>
- Blume, C. (2022). Authentizität. In S.J. Schierholz (Ed.), *Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK)* (Online). Mouton de Gruyter.
- Boersma, P., & Weenink, D. (1992). *PRAAT: Doing phonetics by computer*. (6.1.02) [Computer software]. <http://www.praat.org/>
- Bundgaard-Nielsen, R. L., Best, C. T., Kroos, C., & Tyler, M. D. (2012). Second language learners' vocabulary expansion is associated with improved second language vowel intelligibility. *Applied Psycholinguistics*, 33(3), 643–664.
<https://doi.org/10.1017/S0142716411000518>
- Carey, M. D., Mannell, R. H., & Dunn, P. K. (2011). Does a rater's familiarity with a candidate's pronunciation affect the rating in oral proficiency interviews? *Language Testing*, 28(2), 201–219. <https://journals.sagepub.com/doi/10.1177/0265532210393704>
- Coxhead A., & Bytheway, J. (2014) You will not blink: learning vocabulary using online resources. In D. Nunan & J.C. Richards (Eds.), *Language learning beyond the classroom* (pp. 65–74). Routledge.
- Davies, G., Otto, S. E. & Rüschoff, B. (2013). Historical perspectives on CALL. In M. Thomas, H. Reinders & M. Warschauer (Eds.), *Contemporary computer-assisted language learning* (pp. 19–38). Bloomsbury.
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476–490.
<https://doi.org/10.1017/S026144480800551X>
- Derwing, T. M. & Munro, M. J. (2014). Training native speakers to listen to L2 speech. In J. Levis & A. Moyer (Eds.), *Social dynamics in second language accent* (pp. 219–238). de Gruyter Mouton.
- Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, 27(1–2), 47–55. <https://doi.org/10.1017/S0025100300005417>
- Flege, J., & Bohn, O. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press. <https://doi.org/10.1017/9781108886901.002>
- Gabry, J., & Goodrich, B. (2020). *rstanarm: Bayesian Applied Regression Modeling via Stan*. <https://CRAN.R-project.org/package=rstanarm>

- Galloway, N., & Rose, H. (2014). Using listening journals to raise awareness of global Englishes in ELT. *ELT Journal*, 68(4), 386–396. <https://doi.org/10.1093/elt/ccu021>
- Galloway, N., & Rose, H. (2018). Incorporating Global Englishes into the ELT classroom. *ELT Journal*, 72(1), 3–14. <https://doi.org/10.1093/elt/ccx010>
- Glanz, O., Derix, J., Kaur, R., Schulze-Bonhage, A., Auer, P., Aertsen, A., & Ball, T. (2018). Real-life speech production and perception have a shared premotor-cortical substrate. *Scientific Reports*, 8(1). <https://doi.org/10.1038/s41598-018-26801-x>
- Gronau, Q. F., & Singmann, H. (2021). *Bridgesampling: Bridge sampling for marginal likelihoods and Bayes factors*. <https://github.com/quentingronau/bridgesampling>
- Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, 24(4), 495–522. <https://doi.org/10.1017/S0142716403000250>
- Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, 54(3), 297–326. <https://doi.org/10.1017/S0261444820000658>
- Holliday, A. (2006). Native-speakerism. *ELT Journal*, 60(4), 385–387. <https://doi.org/10.1093/elt/ccl030>
- Hoven, D. (1999). A model for listening and viewing comprehension in multimedia environments. *Language Learning & Technology*, 3(1), 73–90. <http://dx.doi.org/10125/25058>
- Hubbard, P. (2017). Technologies for teaching and learning L2 listening. In C. A. Chapelle & S. Sauro (Eds.), *The Handbook of Technology and Second Language Teaching and Learning* (pp. 93–107). John Wiley & Sons.
- Inceoglu, S. (2021). Language experience and subjective word familiarity on the multimodal perception of non-native speakers' vowels. *Language & Speech*, 1. <https://journals.sagepub.com/doi/10.1177/0023830921998723>
- Ingram, J. C., & Park, S.-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, 25(3), 343–370. <https://doi.org/10.1006/jpho.1997.0048>
- Jacobsen, U. C. (2015). Cosmopolitan sensitivities, vulnerability, and global Englishes. *Language and Intercultural Communication*, 15(4), 459–474. <https://doi.org/10.1080/14708477.2015.1031674>
- Jacquemet, M. (2005). Transidiomatic practices: Language and power in the age of globalization. *Language & Communication*, 25(3), 257–277. <https://doi.org/10.1016/j.langcom.2005.05.001>
- Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford University Press.
- Jolley, K. & Perez, M. (2020). Exploring marginalised communities with online student portfolios using Google Drive and TEDx Talks. In K.-M. Frederiksen, S. Larsen, L. Bradley & S. Thouësny (Vorsitz), CALL for widening participation: Short papers from EUROCALL 2020. (pp. 143-148)
- Jones, M. (2020). Investigating English language teachers' beliefs and stated practices regarding bottom-up processing instruction for listening in L2 English. *Journal of*

- Second Language Teaching & Research*, 8(1), 52–71.
<http://pops.uclan.ac.uk/index.php/jsltr/article/view/590>
- Keating, P. A., & Huffman, M. K. (1984). Vowel variation in Japanese. *Phonetica*, 41(4), 191–207. <https://doi.org/10.1159/000261726>
- Kiczkowiak, M. (2019). Students', Teachers' and recruiters' perception of teaching effectiveness and the importance of nativeness in ELT. *Journal of Second Language Teaching & Research*, 7(1), 1–25.
<https://pops.uclan.ac.uk/index.php/jsltr/article/view/578>
- Kiczkowiak, M. (2021). Pronunciation in course books: English as a lingua franca perspective. *ELT Journal*, 75(1), 55–66. <https://doi.org/10.1093/elt/ccaa068>
- Kozińska, K. (2021). TED talks as resources for the development of listening, speaking and interaction skills in teaching EFL to university students. *Neofilolog*, 56(2), 201–221.
<https://doi.org/10.14746/n.2021.56.2.4>
- Kruschke, J., & Meredith, M. (2021). *BEST: Bayesian Estimation Supersedes the t-Test*.
<https://CRAN.R-project.org/package=BEST>
- Kubota, R. (2015). Inequalities of Englishes, English speakers and languages: A critical perspective on pluralist approaches to English. In R. Tupas (Ed.), *Unequal Englishes: The politics of Englishes today* (pp. 21–41). Palgrave Macmillan.
- Kubota, R. (2021). Global Englishes and teaching culture. In A. F. Selvi & B. Yazan (Eds.), *Language teacher education for global Englishes: A practical resource book* (pp. 135–143). Routledge.
- Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning*, 12(2), 137–156. <https://doi.org/10.1076/call.12.2.137.5722>
- Lotto, A. J., Sato, M., & Diehl, R. L. (2004). Mapping the task for the second language learner: The case of Japanese acquisition of /r/ and /l/. *From Sound to Sense*, 50, 381–386.
- Lowe, R. J., & Kiczkowiak, M. (2016). Native-speakerism and the complexity of personal experience: A duoethnographic study. *Cogent Education*, 3(1), 1264171.
<https://doi.org/10.1080/2331186X.2016.1264171>
- Madarbakus-Ring, N. (2020). Developing graded TED Talks to integrate academic vocabulary into listening lessons for pre-sessional learners. In *The TESOL encyclopedia of English language teaching* (pp. 1–7). John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9781118784235.eelt0990>
- Mathieu, L. (2016). The influence of foreign scripts on the acquisition of a second language phonological contrast. *Second Language Research*, 32(2), 145–170.
<https://doi.org/10.1177/0267658315601882>
- Mayer, R. E. (2021). Cognitive theory of multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 57–72). Cambridge University Press. <https://doi.org/10.1017/9781108894333>
- McCrocklin, S. (2012). Effect of audio vs. video on aural discrimination of vowels. *TESL-EJ*, 16(2), 1–16. <https://www.tesl-ej.org/wordpress/issues/volume16/ej62/ej62a2/>

- Meer, P., Hartmann, J., & Rumlich, D. (2021). Attitudes of German high school students toward different varieties of English. *Applied Linguistics*, *amab046*.
<https://doi.org/10.1093/applin/amab046>
- Montero Perez, M. (2019). Technology-enhanced listening: How does it look and what can we expect? In N. Arnold & L. Ducate (Eds.), *Engaging language learners through CALL: From theory and research to informed practice* (pp. 141–178). Equinox.
- Montero Perez, M., van den Noortgate, W. & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, *41*(3), 720–739.
<https://doi.org/10.1016/j.system.2013.07.013>
- Moodle. (2020). *Moodle*. <https://moodle.com/>
- Mora, J. C., & Fullana, N. (2007). Production and perception of English /i/-/ɪ/ and /æ/-/ʌ/ in a formal setting: Investigating the effects of experience and starting age. *ICPhS XVI*, 1611–1616. <https://tinyurl.com/Mora-J-2007>
- Navarra, J., & Soto-Faraco, S. (2007). Hearing lips in a second language: Visual articulatory information enables the perception of second language sounds. *Psychological Research*, *71*(1), 4–12. <https://doi.org/10.1007/s00426-005-0031-5>
- Nishi, K., & Kewley-Port, D. (2005). Training Japanese listeners to identify American English vowels. *The Journal of the Acoustical Society of America*, *117*(4), 2401–2401.
<https://doi.org/10.1121/1.4786064>
- Nishi, K., & Kewley-Port, D. (2008). Non-native speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research: JSLHR*, *51*(6), 1480–1493. [https://doi.org/10.1044/1092-4388\(2008/07-0109\)](https://doi.org/10.1044/1092-4388(2008/07-0109))
- Norouzian, R., de Miranda, M., & Plonsky, L. (2019). A Bayesian approach to measuring evidence in L2 research: An empirical investigation. *The Modern Language Journal*.
<https://doi.org/10.1111/modl.12543>
- Ockey, G. J., & French, R. (2016). From one to multiple accents on a test of L2 listening comprehension. *Applied Linguistics*, *37*(5), 693–715.
<https://doi.org/10.1093/applin/amu060>
- Piller, I. (2016). *Linguistic diversity and social justice: An introduction to applied sociolinguistics*. Oxford University Press.
- Rącz, P. (2013). *Saliency in sociolinguistics: A quantitative approach*. De Gruyter Mouton.
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing* (4.0.3) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ravelli, L. J. (2018). Multimodal English. In P. Seargeant, A. Hewings, & S. Pihlaja (Eds.), *The Routledge handbook of English language studies* (pp. 434–446). Routledge.
- Rose, H., & Galloway, N. (2019). *Global Englishes for language teaching*. Cambridge University Press. <https://doi.org/10.1017/9781316678343>
- Rubdy, R. (2015). Unequal Englishes: The native speaker, and decolonization in TESOL. In R. Tupas (Ed.), *Unequal Englishes: The politics of Englishes today* (pp. 42–58). Palgrave Macmillan.

- Saito, K., Tran, M., Suzukida, Y., Sun, H., Magne, V., & Ilkan, M. (2019). How do second language listeners perceive the comprehensibility of foreign-accented speech?: Roles of first language profiles, second language proficiency, age, experience, familiarity, and metacognition. *Studies in Second Language Acquisition*, 41(5), 1133–1149. <https://doi.org/10.1017/S0272263119000226>
- Schildhauer, P., Zehne, C., & Schulte, M. (2021). Encountering Global Englishes in the ELT classroom through audio-visual texts: The example of TED talks. In M. Callies, S. Hehner, & P. Meer (Eds.), *Glocalising teaching English as an international language: New perspectives for teaching and teacher education in Germany* (pp. 198-214). Routledge.
- Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Oxford University Press.
- Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3(3), 243–261. <https://doi.org/10.1017/S0142716400001417>
- Shirai, Y. (1990). U-shaped behavior in L2 acquisition. In H. Burmeister & L. P. Rounds (Eds.), *Variability in second language acquisition: Proceedings of the Tenth Meeting of the Second Language Research Forum, Vol. 2* (pp. 685–700).
- Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55(4), 661–699. <https://doi.org/10.1111/j.0023-8333.2005.00320.x>
- Suvorov, R. (2015). The use of eye tracking in research on video-based second language (L2) listening assessment: A comparison of context videos and content videos. *Language Testing*, 32(4), 463–483. <https://doi.org/10.1177/0265532214562099>
- Takaesu, A. (2017). TED Talks as an extensive listening resource for EAP students. In K. Kimura & J. Middlecamp (Eds.), *Asian-focused ELT research and practice: Voices from the far edge* (pp. 108–120). IDP Education.
- Thomopoulos, N. T. (2013). *Essentials of Monte Carlo simulation*. Springer. <https://doi.org/10.1007/978-1-4614-6022-0>
- Trudgill, P., & Hannah, J. (2017). *International English: A guide to varieties of English around the world* (6th edition). Routledge, Taylor & Francis Group.
- Tschirner, E. (2011). Video clips, input processing and language learning. In W. M. Chan, K.N. Chin, M. Nagami, & T. Suthiwan (Eds.), *Media in foreign language teaching and learning* (pp. 25–42). De Gruyter.
- Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second Language Acquisition*, 4(2), 114–136. <https://doi.org/10.1017/S027226310000437X>
- Wu, C.-P. (2020). Implementing TED Talks as authentic videos to improve Taiwanese students' listening comprehension in English language learning. *Arab World English Journal*, 6, 24–37. <https://doi.org/10.24093/awej/call6.2>

Appendix 1: List of TED Talks used

- Adegbeye, O. (2017). *Who belongs in a city?*
https://www.ted.com/talks/olutimehin_adegbeye_who_belongs_in_a_city

- Belle, R. A. (2021) *The emotions behind your money habits*.
https://www.ted.com/talks/robert_a_belle_the_emotions_behind_your_money_habits?language=en
- Cekic, Ö. (2018). *Why I have coffee with people who send me hate mail*.
https://www.ted.com/talks/ozlem_cekic_why_i_have_coffee_with_people_who_send_me_hate_mail
- Francis, A. (2019). *How to get everyone to care about a green economy*.
https://www.ted.com/talks/angela_francis_how_to_get_everyone_to_care_about_a_green_economy
- Gibson, E. (2019). *The true cost of financial dependence*.
https://www.ted.com/talks/estelle_gibson_the_true_cost_of_financial_dependence
- Hegde, A. (2021). *The life-saving tech helping mothers make healthy decisions*.
https://www.ted.com/talks/aparna_hegde_the_life_saving_tech_helping_mothers_make_healthy_decisions
- Howell, E. (2018). How we can improve maternal healthcare -- before, during and after pregnancy.
https://www.ted.com/talks/elizabeth_howell_how_we_can_improve_maternal_healthcare_before_during_and_after_pregnancy
- Jun, M. (2021). An interactive map to track (and end) pollution in China.
https://www.ted.com/talks/ma_jun_an_interactive_map_to_track_and_end_pollution_in_china
- Pearlman, E. (2019). How to lead a conversation between people who disagree.
https://www.ted.com/talks/eve_pearlman_how_to_lead_a_conversation_between_people_who_disagree
- Speck, J. (2013). The walkable city. https://www.ted.com/talks/jeff_speck_the_walkable_city

Appendix 2: List of R packages used

- Gabry, J., & Goodrich, B. (2020). *rstanarm: Bayesian Applied Regression Modeling via Stan*.
<https://CRAN.R-project.org/package=rstanarm>
- Gronau, Q. F., & Singmann, H. (2021). *bridgesampling: Bridge Sampling for Marginal Likelihoods and Bayes Factors*. <https://github.com/quentingronau/bridgesampling>
- Guo, J., Gabry, J., Goodrich, B., & Weber, S. (2020). *rstan: R Interface to Stan*. <https://mc-stan.org/rstan/>
- Kruschke, J., & Meredith, M. (2021). *BEST: Bayesian Estimation Supersedes the t-Test*.
<https://CRAN.R-project.org/package=BEST>
- Simonsohn, U., & Gruson, H. (2021). *groundhog: The Simplest Solution to Version-Control for CRAN Packages*. <https://CRAN.R-project.org/package=groundhog>
- Spinu, V., Grolemund, G., & Wickham, H. (2021). *lubridate: Make Dealing with Dates a Little Easier*. <https://CRAN.R-project.org/package=lubridate>
- Wickham, H. (2021). *tidyverse: Easily Install and Load the Tidyverse*. <https://CRAN.R-project.org/package=tidyverse>

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>

Wilke, C. O. (2020). *cowplot: Streamlined Plot Theme and Plot Annotations for ggplot2*. <https://wilkelab.org/cowplot/>

Appendix 3: List of words tested

All recordings are available to download from the project's Open Science Framework page. https://osf.io/xb9ce/?view_only=e420caa1444a42aba486a129e71920a0

bad, ban, board, born, bun, burn,

call, can, cork, cup

dad

fawn, fun

girl

ham, heard, hug, hut,

kern,

lash, luck

nurse

pause, pearl, pork

sad, shut, sword,

torn, turn

yearn

zap

Copyright of articles rests with the authors. Please cite TESL-EJ appropriately.