Examining the Segmental and Suprasegmental Correlates of the IELTS Pronunciation Scale

Author: Yui Suzukida

Birkbeck, University of London

1   **Examining the Segmental and Suprasegmental Correlates of the IELTS**

2   **Pronunciation Scale**

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20 **Abstract**

21   The current study identified the crucial segmentals and suprasegmentals for

22 developing IELTS pronunciation syllabus and explored in depth the extent to which trained

23 raters actually follow the IELTS rubric to distinguish different levels of candidates in order

24 to suggest which pronunciation features test takers should selectively work on for ultimate

25 improvement of overall test results. This study additionally investigated the practicality of

26 improving IELTS pronunciation skills.

27   Forty speech samples of native Japanese speakers were first rated by seven trained

28 judges based on the publically available IELTS pronunciation rubric and degree of

29 comprehensibility, then the same samples were objectively coded by the researcher

30 according to eight pronunciation measures and six problematic segmental groups for

31 Japanese learners.

32   The results of the correlation analysis suggested that more errors in productions of

33 major segmentals (/l, ɹ, ð, q, v/), syllable and word stress lower the raters' evaluations on

34 IELTS pronunciation marking. Confidence intervals revealed that accurate production of

35 major segmentals (/l, ɹ, ð, q, v/), secondary segmentals (/æ, ʌ, f/), syllables, and diphthongs

36 are important for beginner group (Band scores 4–5), while accurate production of word

37 stress as well as the minor segmentals (/w, ŋ, h, n, p, t, k/, Contractions) are crucial for

38 intermediate group (Band score 6) to attain advanced level (Band scores 7–8). Finally, the

39 result of the correlation analysis yielded significantly high correlation of IELTS

40 pronunciation rating with comprehensibility judgements, supporting the high applicability of

41 the IELTS pronunciation skill to the success in L2 communication in real-life context.

42

43

Table of Contents

104

105

106

107

108

109

110

111

112

113

114

115

116

117

119

Tables

129

Figures

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

Chapter 1: Introduction

169

170   **1.1 Background**

171

172   In a long history of controversy among second language (L2) research on the

173   development of pronunciation teaching, one of the central focuses is what model to follow

174   for instruction (Walker, 2010). Most recently, due to the expansion of variety and use of

175   English following increasing globalisation, the standard of pronunciation instruction has

176   been subject to changes from pursuing established regional varieties of English such as

177   British English Received Pronunciation to achieving mutual intelligibility for wider

178   communication in a global context (Brinton, 2012). Despite such transition in the standards

179   however, most language instructions and pronunciation syllabi still hold implicit emphasis

180   on the native speaker model and teachers' intuitive decisions based on their norm of native-

181   like pronunciation (Murphy, 2014), indicating the pressing need for fostering an awareness

182   and development of the legitimate instructional model instead of mastering native-like

183   ability.

184   Beside the communicative success in L2 oral communication in a global context,

185   what educators need to be aware of when teaching L2 pronunciation is the needs of

186   individual learner. In effect, quite a lot of learners of English who pursue academic careers

187   in English-medium universities are eager to learn English pronunciation to obtain a

188   certificate in their English exams. This is because the certificate is used as a tangible proof of

189   one's sufficient L2 proficiency for meeting the language requirements set by those

190   institutions for admission purposes. In the context of the UK, International English

191   Language Testing System (IELTS) is a dominant and recommended test to prove one's

192   proficiency level when applying for UK universities (British Council, n.d.-b). Since Band 7

193   on 9-point scale is commonly regarded as a benchmark to be accepted by these universities,

194   candidates of the IELTS are in need of instructions for acquiring sufficient skills in every

195   four skill (speaking, reading, writing, and reading).

196

197   **1.2 Rationale**

198

199   Importantly, it should be mentioned that there is already a plethora of evidence that

200   shows the importance of pronunciation skills in order for one to be successful in the

201   speaking section of the English proficiency test. For instance, Iwashita, Brown, McNamara

202   and O'Hagan (2008) speculated that pronunciation errors could be an initial determining

203   factor for examiners to make an overall proficiency judgment based on their finding that

204   some learners, who have high vocabulary richness and poor pronunciation skill were

205   assigned lower scores by raters. Thus, pronunciation is a crucial skill, which could

206   negatively impact on overall rating and should be a first challenge for learners to overcome

207   in order to make their speech assessable by examiners. Notably, in the case of IELTS

208   pronunciation assessment, interplay of examiners' rating process and their use of the scales

209   have remained underresearched (see Isaacs, Trofimovich, Yu & Chereau, 2015). Given the

210   minor status and unsystematic organization of pronunciation instruction within the speaking

211   syllabus in general (Breitkreutz, Derwing & Rossiter, 2001) and a pressing need from

212   learners to develop test-taking strategies for speaking tests (Issitt, 2008), it is reasonable to

213   examine the legitimate pronunciation instructions and develop a tailor-made pronunciation

214   syllabus of IELTS for test-takers who seek academic success in their careers. The

215   development of such a syllabus would be especially helpful for native speakers of Japanese

216   (NJs), who often receive limited amount of instructions on proper pronunciation in the

217   classroom setting and need different instructions specialised for taking the IELTS test in

218   order to study in UK universities. In fact, researchers claimed that Japanese learners face

219   difficulties in articulating English in terms of various dimensions ranging from segmentals

220   (Nishi & Kewley-Port, 2007; Riney & Takagi, 1999; Sekiyama & Tohkura, 1993) to

221   suprasegmental levels (Ohata, 2004; Tsujimura, 2013) due to the great difference in phonetic

222   system from that of English.

223

224   **1.3 Aim and Structure of the Current Study**

225

226   Therefore, in order to move beyond the current L2 pronunciation research and cast

227   light on the development of pronunciation syllabus for test-taking purpose, the current study

228   aim at developing strategic pronunciation teaching for IELTS test-takers, and in-depth

229   understanding of the constructs to be measured in IELTS pronunciation skill. Also, the

230   present study examines the adaptability of the IELTS pronunciation skills in real-life

231   communication by setting the degree of correlation with comprehensibility as a barometer.

232   This thesis is organised in five parts: the first section overviews previous literature,

233   which illustrates the underresearched respects of pronunciation teaching and generates the

234   research questions of this study. Then the second section addresses the detailed account of

235   the methodological approach and research design. The subsequent section begins with the

236   presentation of the results, followed by the discussion with respect to the three research

237   questions. Lastly, the findings of the study are summarised with research implications and

238   limitations.

239

240

241

242

Chapter 2: Literature Review

244

245 **2.1 Introduction**

246

247   This chapter discusses the L2 oral proficiency model for language teaching/learning

248 and the underlying issue of the adoptability in applying research outcomes to the context of

249 the IELTS pronunciation skill development. The description of the conceptual change in L2

250 speaking proficiency principles (i.e., comprehensibility as a feasible goal) is introduced with

251 empirical evidence, which is followed by the overview of the key components of the highly

252 comprehensible pronunciation in past research findings. Subsequent section raises issues in

253 the modelling of comprehensible pronunciation teaching in the IELTS pronunciation

254 assessment.

255

256 **2.2 Teaching L2 Pronunciation**

257

258   **2.2.1 Paradigm shift in L2 oral proficiency.** L2 pronunciation research had

259 experienced a major paradigm shift in its pedagogical practice/goal from *nativelike*

260 pronunciation (i.e., accent-free speech) to intelligible pronunciation (i.e., accented but

261 understandable pronunciation; Derwing & Munro, 2005; Levis, 2005). According to Levis

262 (2005), this shift involved two contradictory principles: *Nativeness principle* which is

263 defined as the instructions for the eradication of a foreign accent whereas *intelligible*

264 *principle*'s main focus is described as the instructions for the production of understandable

265 speech with a foreign accent. Although there are handful of exceptional cases (see Derwing,

266 2003; Kang, 2010; Scales, Wennerstrom, Richard & Wu, 2006; Timmis, 2002; Tokumoto &

267 Shibata, 2011), researchers have found that unattainability of accent-free speech is due to

268   various factors such as learners' language background (Flege, 2003), age of acquisition

269   (Flege, Munro, & MacKay, 1995), quality and quantity of input (Fledge, Yeni-Komshian, &

270   Liu, 1999), attitude, motivation, and aptitude toward L2 speech learning (Moyer, 1999).

271   Therefore, foreign accented but intelligible pronunciation is the realistic goal for L2 learners

272   hoping to achieve international communicative success (e.g., Derwing & Munro, 2009;

273   Tutor et al., 2016a; Tutor et al., 2016b; Tutor et al., 2016c).

274   **2.2.2 Defining comprehensibility.** With the shift of paradigm, enormous amounts of

275   research have been carried out to deconstruct the components of intelligible and

276   unintelligible pronunciation though listener-judgment of *accentedness* and *comprehensibility*

277   (for review of the various approach, see Derwing & Munro, 2005). Accentedness refers to

278   the listeners' perception of how close the speaker's language is to the speech patterns of the

279   target-language community (Munro & Derwing, 1999), thus best described as *linguistic*

280   *nativelikeness* (Tutor et al., 2015c), while comprehensibility is frequently used in L2

281   pronunciation research as the synonym of intelligibility (Levis, 2005). Although Isaacs

282   (2008) pointed out that the confusion in the use of intelligibility between each study is due to

283   its manifold definition and interpretation of the term (for more discussion of the term, see

284   Isaacs, 2008), a rather clear account has been provided by Munro & Derwing (1999).

285   According to their explanation, intelligibility and comprehensibility both indicate the

286   listener's ability to understand L2 speech in the broader meaning (see Levis, 2005), yet, in a

287   narrower sense, they are different in the way they measure the listener's understanding:

288   intelligibility measures the amount of understanding via orthographical transcribing whereas

289   comprehensibility measures the degree of understanding through a scalar rating. Thus, the

290   distinction of the two terms seems to be made by the way of operationalising the

291   measurement (Isaacs et al., 2015). This study, therefore, defines accentedness as the listeners'

292   perceptual degree of nativeness of the L2 speech, comprehensibility as the listeners' degrees

293   of understanding of L2 speech based on rating, and intelligibility as, in a broader sense,

294   listeners' ability to understand L2 speech. Assuming that learners can communicate

295   successfully with an accented L2 (Derwing & Munro, 2009), researchers have been

296   attempting to reveal the unique and shared linguistic factors that underlie comprehensibility

297   and accentedness for developing practical L2 pronunciation pedagogy.

298   **2.2.3 Role of pronunciation in comprehensibility.** To date, while the

299   distinctiveness of comprehensibility and accentedness were argued (e.g., Jenkins 2000; Kang,

300   Rubin & Pickering, 2010; Munro & Derwing, 1995), research evidence supports that

301   pronunciation features both segmental and prosodic aspects impact on both accentedness and

302   comprehensibility judgments. Accentedness seems to be strongly associated with segmental

303   accuracy, temporal measures, syllable duration, stress, and pitch range (Anderson-Hsieh,

304   Johnson & Koehler, 1992; Winters & O'Brien, 2013), while comprehensibility is linked to

305   segmental sounds in stressed syllables (Zielinski, 2008), word stress (Field, 2005), primary

306   stress in sentences (Hahn, 2004), tone choice (Pickering, 2001; Wennerstom, 2001) as well

307   as lexicogrammatical accuracy (Munro & Derwing 1995; Varonis & Gass 1982).

308   While these studies seem to have a bias towards the negative impact of non-

309   segmental aspects of speech, without a shadow of a doubt it is clear that some segmental

310   difficulties should be given attention as well. One of the segmental-focused studies on

311   comprehensibility, for example, was conducted by Munro and Derwing (2006) who selected

312   the segmentals from Functional Load (FL) theory, which is a ranking list of segmental

313   contrasts based on their importance in English pronunciation, developed by Brown (1991)

314   and Catford (1987). The finding revealed that high FL (/l, ʃ, n, s, d/) errors were rated worse

315   than that of low FL (/ð, θ/) to greater extent on both accentedness and comprehensibility

316   (Munro & Derwing, 2006). This indicates that not all but certain segmentals appear to be

317   crucial for comprehensibility, underpinning the essential contribution of segmentals on

318   comprehensibility.

319   These comprehensibility studies were more extensively examined by Isaacs and

320   Trofimovich (2012), and Trofimovich and Isaacs (2012), who employed various ranges of

321   measures including pronunciation, fluency, lexis, grammar, and discourse structure for

322   linguistic coding. For instance, Trofimovich and Isaacs (2012) found that comprehensibility

323   and accentedness were considerably linked to pronunciation measures such as word stress

324   and rhythm, but comprehensibility was also influenced by grammatical accuracy and lexical

325   type frequency and accentedness judgments were uniquely tied to the listeners' perceptual

326   salience of vowels and consonants errors (Trofimovich & Isaacs, 2012). However, while a

327   follow-up study by Tutor et al. (2015c) confirmed the impact of word stress errors on both

328   accentedness and comprehensibility, they also revealed that segmental errors not only affect

329   accentedness but also comprehensibility judgement. Furthermore, the result indicated that

330   the overall rating score of comprehensibility were higher than that of accentedness,

331   supporting the claim that one's accented L2 speech can be highly comprehensible (Derwing

332   & Munro, 2009; Jenkins, 2000; Munro & Derwing, 1995).

333   **2.2.4 First language influence.** Yet there is a general consensus that one's first

334   language (L1) can impact on L2 development and this topic has been widely discussed in the

335   L2 phonological development literatures with multiple learning models (e.g., Darcy,

336   Dekydtspotter, Sprouse, Glover, Kaden, McGuire & Scott, 2012; Eckman, 2004; Eckman &

337   Iverson, 2013; Escudero & Boersma, 2004), little attention has been paid to how speaker's

338   L1 come into play in listener's judgment of comprehensibility. One of the predominant

339   approaches used in the L2 development research is Lado's (1957) contrastive analysis

340   hypothesis (CAH), which gives an account of L2 learners' phonological development in

341   terms of similarities and differences between speaker's L1 and L2 phonologies. On one hand,

342   differences of speaker's L1 and L2 are emphasized to describe the pattern of L2

343   development (e.g., Lado, 1957), on the other hand, some researchers focus on the similarities

344   of these two languages (Oller & Ziahosseiny, 1970; Young-Scholten, 1985), however, both

345   lines of the research are devoted to giving the account of the difficulty L2 learners face (for

346   reviews, see Eckman, 2004; Major, 2008). Yet still underdeveloped, this array of studies

347   demonstrated that learners' learning difficulties and error types in L2 pronunciation may

348   vary depending on their L1 backgrounds (Eckman & Iverson, 2013).

349   Pertaining to listener-based comprehensibility studies, however, the L1 variable has

350   not receive sufficient attention except for few studies such as Tutor et al. (2015a). They

351   looked at how L1differences affect the correlation between errors in linguistic features and

352   degree of comprehensibility. Although no strong association was found with Farsi speakers,

353   their results showed that comprehensibility was linked to pronunciation variables with

354   Chinese speakers, lexicogrammar with Hindi speakers, both pronunciation and

355   lexicogrammar variables with Romance speakers, indicating that L1 background certainly

356   bears relevancy to comprehensibility. Thus, further expansion and refinement of the research

357   is required to find which features are uniquely tied to the specific L1 group's L2

358   comprehensibility.

359   **2.2.5 Setting priority in pronunciation teaching.** The range of studies discussed

360   above have empirically proven that pronunciation features especially word stress, segmentals

361   and rhythm seem to be the overlapping components of comprehensible speech and native-

362   like speech. However, regarding the development of pronunciation syllabi for teaching, these

363   research outcomes are sparse especially with respect to segmentals. While non-segmental

364   features are relatively clear on the importance of certain non-segmental features over others

365   (e.g., word stress), further clarification needs to be done on segmentals. This is because they

366   often treated as a homogeneous measurement category albeit they are an aggregation of the

367   individual sounds except for one exception of Munro and Derwing's (2006) use of

368   Functional Load theory. Yet this study confirmed that certain segmentals should take

369   precedence over other segmentals due to their different degree of impact on

370   comprehensibility, fine-grained research on segmental classification is fairly limited. Beside

371   the Functional Load theory (Brown, 1991; Catford, 1987), Jenkins (2002), from the world

372   English paradigm, sought to find the crucial pronunciation features and its teaching priorities

373   to attain global intelligibility between non-native speakers of English (NNEs). Her

374   pronunciation syllabus (i.e., *Lingua Franca Core*) is based on the observation of the

375   breakdowns of interaction between NNEs, and comprise of *core* and *non-core* which

376   separate crucial features especially individual segmentals from the peripheral features.

377   Fairly recently, Tutor (2014) developed an exhaustive list of segmentals and

378   suprasegmentals to be prioritized for native Japanese learners of English. Based on the cross-

379   linguistic analysis between English and Japanese, 25 problematic pronunciation features

380   including specific segmentals were carefully selected and rated on importance by 120

381   experienced teachers (NEs and NJs). The results revealed that pronunciation needs to be

382   taught in the following order: major segmentals (/l, ɹ, ð, q, v/), L1 effect at syllable levels

383   (cognates, syllabification), assimilation (/si, ʃi, ti/), stress/ intonation (sentence/ lexical stress,

384   intonation), secondary segmentals (/æ, ʌ, f/), diphthongs (/aʊ, aɪ, oʊ, ɔɪ, eɪ/)□□minor

385   segmentals (/p, t, k, w, n, ŋ, h/) and contractions, and fluency problems (Tutor, 2014).

386   Compared to the Lingua Franca Core (Jenkins, 2002) practice which considers NNEs of

387   various language backgrounds and deals with universally essential features for NNEs, the

388   uniqueness of Tutor's (2014) study is its focus on a single L1 group (i.e., Japanese) with its

389   theoretical basis from L2 phonological development theory (i.e., contrastive analysis

390   hypothesis). Nevertheless, research on prioritisation of certain segmentals is still

391   unsatisfactory to identify universal and L1 specific segmental features to develop

392   pronunciation syllabus for improving one's comprehensibility.

393

2.3 Issues in Realising Pronunciation Instructions for the IELTS Pronunciation Skill

395

Although the research on constructs of comprehensibility certainly informs

instructional agenda of L2 pronunciation, practitioners need to consider how such empirical

evidence can be applied to the specific contexts, where learners' success is beyond the

accomplishment of L2 communication. In the academic context, L2 learners, who pursue

academic careers in English-medium universities, need to obtain the skill to meet the

standard measured through the English proficiency tests as a part of their application to the

university (Feast, 2002).

Among UK universities, for instance, International English Language Testing System

(IELTS) is the dominant and preferred high-stake English proficiency tests for assessing

applicants for admission purposes. The IELTS score, which consists of the discrete marks of

four disciplines (listening, writing, reading, and speaking) and overall score, is regarded as a

predictor of the applicants' proficiency levels to be successful in their courses (Hayatt &

Brooks, 2009; Hayes & Read, 2004; Issitt, 2008). Upon selection, most higher education

providers set the minimum requirements of the IELTS score (for more detail of the

minimum Band scores for application, see British Council, n.d.-c). The IELTS assesses the

test-takers pronunciation skill as one of four subcomponents of their speaking skill (i.e.,

fluency, grammatical accuracy, lexical richness) together with the other three skills (i.e.,

listening, writing, reading).

With respect to IELTS pronunciation skill development, there are at least two issues

to consider for operationalising comprehensibility practice: (a) the inconsistency of the

measuring construct, (b) vagueness in the rating descriptor. Firstly, the construct (conceptual

model) being measured in IELTS pronunciation skill is inconsistent (Isaacs et al., 2015) thus

418 the appropriateness of the adaptation of comprehensibility practice to IELTS pronunciation

419 skill is uncertain. Isaacs et al. (2015) examined the pronunciation scales consulted by

420 accredited IELTS examiners (the scales are different from publically available version), and

421 found that the scale includes a mixture of three constructs (intelligibility, comprehensibility,

422 and accentedness), which seems to be a common shortcoming among L2 oral proficiency

423 tests (Harding, 2013; Isaacs & Trofimovich, 2012). The actual criteria for the score

424 judgement are provided by Seedhouse, Harris, Naeb, and Üstünel (2014): "Pronunciation

425 refers to the capacity to produce comprehensible speech in fulfilling the speaking test

426 requirements. The key indicators will be the amount of strain caused to the listener, the

427 amount of unintelligible speech and the noticeability of L1 influence" (p. 5). According to

428 the definitions discussed above, it can be speculated that the terms "comprehensible speech"

429 and "strain caused to the listener" refer to comprehensibility, and "unintelligible speech"

430 corresponds to intelligibility as a synonym of comprehensibility, since IELTS pronunciation

431 assessment involves examiners' subjective scalar ratings, which is the same measuring

432 system as comprehensibility (Isaacs et al., 2015), indicating that their construct is built on a

433 comprehensibility paradigm. However, "the noticeability of L1 influence" is associated with

434 accentedness. Thus, from the researchers' perspective, it is not clear to what extent the

435 degree of accentedness comes into play in the raters' judgement, as empirical evidence

436 supports that a L2 speaker with an accent can be highly comprehensible (Isaacs et al., 2015).

437   On top of the examiner's assessment criteria, the ambiguity also exists in key

438 indicators of the public rubric (see Appendix A). Comprehensibility-associated indicators are

439 observed in Band 9 ("effortless to understand"), Band 8 ("easy to understand throughout"),

440 Band 6 ("generally be understood throughout"), Band 4 ("cause some difficulty for the

441 listener"), and Band 2 ("unintelligible"), while accentedness-related indices do exist in Band

442 8 ("L1accent has minimal effect"), and in Bands 6 and 4 ("mispronunciation"). Such

443   conflation may cause teachers and learners difficulty in understanding what model to follow

444   for improving IELTS pronunciation skill (i.e., nativelike or comprehensibility model).

445   The second issues underlying the operation of IELTS pronunciation instruction is

446   the vagueness of the distinction between each band due to the lack of information in

447   pronunciation features illustrated in the public version of score indicators. The vagueness

448   issue is two folds: vagueness in the description of pronunciation features in Bands 4, 6, 8,

449   and 9, and vagueness in the description of Bands 3, 5, and 7. Regarding the first vagueness,

450   with the exception of Band 2 ("often unintelligible") and Band 1 ("no communication

451   possible, no rateable language"), the differentiation of the expected pronunciation features to

452   be performed in Bands 2, 4, 6, 8 and 9 is crude: "a full range" in Band 9, "a full range" in

453   Band 8, "a range" in Band 6, "a limited range" in Band 4 respectively. This vagueness of the

454   features indicates that further specification is needed for clear understanding of what

455   components pronunciation features contain (i.e., segmentals and suprasegmental features)

456   for teachers and examiners (Isaacs et al., 2015). The second vagueness is the

457   underspecification of Bands 3, 5, and 7 compared to that of relatively detailed explanation of

458   Bands 2, 4, 6, 8 and 9. As Isaacs et al. (2015) summarized, these three bands are briefly

459   defined by referring to the descriptions of the band description immediately below and

460   above. For instance, Band 7 is defined as "shows all the positive features of Band 6 and

461   some, but not all, of the positive features of Band 8". This vagueness of definition in in-

462   between bands (Yates, Zielinski, & Pryor, 2011, p. 34), again creates greater confusion for

463   examiners in assigning the score as well as teachers and test-takers to clarify what they need

464   to aim for (Isaacs, et al., 2015). For Band 7, this vagueness is especially problematic for

465   international students who seek a degree in English-medium universities since Band 7 is

466   often used as a threshold for the admission purposes in the universities. Therefore, although

467   the pronunciation rubric has been revised to a 9-point scale (for the detail of the history of

468   revision, see DeVelle, 2008) due to its coarse classification of the old 4-point based old scale

469   where (Brown, 2006), the indicators still lack of its detailed discrimination between bands

470   and components of the pronunciation features they assess.

471   Although direct contribution to the IELTS pronunciation syllabus has not been made,

472   a study by Isaacs and her colleagues' (2015) provides insights into the two issues discussed

473   above. They examined the linguistic correlates, which are attended to by the accredited

474   IELTS examiners when they discriminate between scores in the pronunciation rating as well

475   as how comprehensibility is modelled in the actual rating. For their study, 80 pre-rated

476   speeches of the IELTS candidate from various L1 background such as Chinese, Arabic,

477   Tagalog, Spanish, Thai, Kannada were re-rated by eight IELTS raters in terms of seven

478   linguistics measures (i.e., grammatical accuracy, lexical richness, speech chunking, vowel

479   and consonant errors, word stress, intonation, and speech rate), comprehensibility and IELTS

480   pronunciation rubric. The pronunciation rating was strongly correlated with grammatical

481   accuracy, lexical richness and speech chunking whilst most of the pronunciation features

482   such as vowel and consonant errors, word stress, intonation and speech rate had a low

483   correlation with the rating. Notably, they also reported that comprehensibility showed the

484   least correlation with the IELTS pronunciation rating. These results indicated that the IELTS

485   pronunciation skill is not just about comprehensibility but other factors were taken into

486   account for the judgement. Also, pronunciation features had little impact on the rating but

487   were influenced by other factors such as lexicogrammar. With regard to the score

488   discriminating factors, all the measures contributed to the discrimination between Bands 6

489   and 7. In addition, speech rate and lexical richness also distinguished Bands 5 and 6, but

490   other measures were not strong determinants of separating Bands 5 and 6 or Bands 7 and 8.

491   Their findings shed light on the issue of rubric vagueness and inconsistency of the measuring

492   construct in IELTS pronunciation assessment by revealing linguistic factors involving the

493   examiners' actual rating process and discrimination of the bands. Their investigation is the

494   very first study, which looked at IELTS pronunciation assessment and explored the

495   phonological features, which differentiate the pronunciation scores. Therefore, setting this

496   study as a starting point, further details need to be researched in terms of the contribution of

497   pronunciation features such as the weight of individual segmentals. In addition, it is worth

498   examining how a learner's L1 background comes into play on the pronunciation rating since

499   research evidence on L2 pronunciation development and comprehensibility discussed above

500   have supported the impact of the L1 on L2 production (e.g., Tutor et al., 2015b).

501

502   **2.4 Summary**

503

504   This section discussed the literature regarding the key concepts and findings involved

505   in this present research. First, the paradigm shift in L2 pronunciation model was introduced

506   by an overviewing of the key literatures. Due to the ample amount of research evidence of

507   unfeasibility in attaining nativelike pronunciation, researchers have proposed that

508   comprehensibility is a realistic goal for L2 learners, and explored the linguistic factors,

509   which contribute to L2 comprehensibility. Followed by the defining of three core terms in

510   this study (Intelligibility, comprehensibility, and accentedness), the essential constructs of

511   comprehensibility and accentedness have been described through comparison of relevant

512   studies. Through the examination of the research findings, it was confirmed that

513   pronunciation features not only impact accentedness but also on comprehensibility. Further,

514   based on the L2 pronunciation development literature, the influence of learner's L1 on

515   comprehensibility was presented in the latest findings of Tutor et al. (2015b). Then, the lack

516   of the detail in the selection of segmentals for the development of a pronunciation syllabus

517   towards comprehensibility was pointed out, followed by the introduction of available

518 research evidence, which attempt to describe the priorities of certain segmentals. Lastly, the

519 relevancy of developing the syllabus for IELTS pronunciation teaching was emphasised and

520 the two issues in adopting the comprehensibility research outcome to the IELTS context was

521 pointed out with the findings of a key study on linguistic correlates of the IELTS

522 pronunciation rating.

523

524 **2.5 The Current Study**

525

526   As we reviewed in the previous section, the scope of the pedagogical research of

527 pronunciation does not cover the development of instructional guidelines for testing contexts

528 such as the IELTS pronunciation assessment. Also, the relevant evidence of the IELTS

529 pronunciation rating from the language testing research side is limited to one study done by

530 Isaacs et al. (2015). Therefore, in order to establish a strategic guideline for the IELTS

531 pronunciation assessment, further research is required with an in-depth examination of the

532 contribution of individual segmentals and test-takers' L1 backgrounds to the rating results.

533   Thus, to advance the current development of the pronunciation teaching and deepen

534 the latest understanding of the key constructs measured by the IELTS pronunciation ratings,

535 the present study set its goals to identify crucial pronunciation features and segmentals

536 which link to the IELTS pronunciation rating itself and discrimination of the band scores by

537 focusing on one L1 group (i.e., Japanese adult learners). This study also explores the

538 relationship between comprehensibility and IELTS pronunciation skill to examine the

539 practicality of the IELTS pronunciation skill for real life communication success. The

540 research questions are as follows:

541

542   1. Which pronunciation aspects are priorities for teaching the IELTS pronunciation

543   skill to native Japanese learners of English in order to improving their score?

544

545   2. What pronunciation aspects most distinguish different proficiency levels of NJ

546   learners of English in IELTS pronunciation rating?

547

548   3. Is the IELTS pronunciation skill useful for the learner's real life communication

549   success?

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566   **Chapter 3: Methodology**

567

568   **3.1 Introduction**

569   In this chapter, the methodological approach and research design used to arrive at the

570   answers to the research questions are introduced together with the discussion of its suitability

571   for the current research. This chapter also covers data collection procedure and how they are

572   analysed.

573   The first section gives an account of a quantitative approach employed in the current

574   study, this is then followed by the overview of the research design, which contains the

575   illustration and justification of the methods adopted to obtain the valid data for examining

576   the research questions. The subsequent section involves the description of the data collection

577   procedure including the participants, instruments and data collection step, which lead into

578   the elucidation of data analysis.

579

580   **3.2 Methodological approach**

581

582   In line with the series of research on L2 prosodic features and development (e.g.,

583   Tutor et al, 2015a; Tutor et al., 2016b), the current study employs a quantitative approach to

584   seek answers for the three research questions (cf. Chapter 2). The nature of quantitative

585   approach is to deal numerical data and enables researchers to run descriptive analysis and

586   inferential statistics (Lazaraton, 2005; Loewen & Plonsky, 2016). Since the current study

587   seeks to identify the numbers of errors L2 speakers make, how many errors different

588   proficiency groups make (i.e., mean of the errors and possible error ranges), how high and

589   low the speech is rated based on the 9-point scale (i.e., average scores) for the analysis,

590   quantitative approach is regarded as best suited. For data collection**,** in order to analyse the

591 relationship between assigned IELTS pronunciation score and the number and types of

592 errors in the speech, speech rating by native judges and phonological error coding are used.

593 These two methods are predominant combination in the empirical research (e.g., Isaacs &

594 Trofimovich, 2012; Trofimovich & Isaacs, 2012).

595   Regarding listener-based speech rating, it has been extensively used for L2 speech

596 evaluation. While rating consistency has been a focus of the debate among number of studies

597 (e.g., Calloway, 1980; Thompson, 1991, Rossiter, 2009), relatively high degrees of

598 reliability have been proven by rater-judgement study of segmental judgement

599 (Cunningham-Anderson & Engstrand, 1989; Piske, MacKay, & Fledge, 2001), and

600 comprehensibility and accentedness judgements (Isaacs & Thomson, 2013), suggesting that

601 scale-based listener judgements hold sufficient reliability for the current study. As for an

602 ethical issue of participant recruitment for rating, present study followed ethical research

603 procedure set by author's university, and obtained an ethical approval from the department

604 ethics officer. Phonological coding or linguistic coding in broader sense, have been directed

605 either auditory measure to count the frequency of the specific phenomenon or coding

606 software such as speech coding software *Praat* (Boersma & Weenink, 2010) or lexical tutor

607 (Cobb, n.b.) for measuring lexical aspects. While some studies such as Trofimovich and

608 Isaacs (2012) used both auditory and analysis software, present study only employed

609 auditory measure for error frequency counting in order to focus on errors that are detectable

610 to teachers.

611

612 **3.3 Instruments**

613

614   **3.3.1 Participants.** The participant (henceforth, raters) for speech rating in the

615 present study included seven teachers (M_{age} = 36 years, range = 29–56) recruited via email in

616 London. They speak English as their L1 and were borne and/or raised in English-speaking

617 home environment with at least one English-speaking parent except for one rater whose

618 parents' L1 were Urdu. They reported that English is the major medium of the

619 communication (M = 96 %, range = 90–100). In order to minimise the inconsistency and

620 inaccuracy in evaluating the speech, the current study controlled two rater variables via

621 language background questionnaire (see Appendix B): (a) rater's familiarity with non-native

622 speakers L2 (see Isaacs & Trofimovich, 2012 for detailed discussion of listener factor), and

623 (b) raters' linguistic and teaching experience (for the high reliability and accuracy in

624 pronunciation rating of linguistic and teaching experience, see Tutor et al., 2015c). With

625 respect to familiarity of Japanese-accented English, the raters reported high familiarity with

626 Japanese-accented English (M = 5.14, range = 4–6) on 6-point likert scale (1 = *not at all*, 6 =

627 *very much*) with frequent contact with NJs (M = 5, *range* = 4–6) on 6-point likert scale (1 =

628 *very infrequent*, 6 = *very often*). Four out of seven have experience of visiting Japan (two

629 weeks, one month, six month, and two years) and two of the four raters took short-term

630 Japanese language courses. Thus, all the raters were regarded as consistent in their

631 familiarities with Japanese accented English. Considering their linguistic knowledge and

632 teaching background, all the raters responded that they have teaching experience in

633 EFL/ESL contexts (M = 9.9 years, *range* = 3-25), and have extensive linguistic/phonological

634 knowledge obtained through either enrolled in or completed their master degree in Applied

635 Linguistics, TESOL, or TEFL courses in English-medium university. Therefore, the seven

636 raters are regarded as relatively homogenous in their phonological and linguistic knowledge

637 as well as their teaching experience.

638   **3.3.2 Speech data.** The present study used 40 speech data from Tutor's (2011)

639 unpublished corpus, which contains more than 200 audio data collection of Japanese learners

640 who completed various speaking tasks in Japan and Canada. For the sake of the

641   investigation, careful selection had been made to include various proficiency levels of the

642   speakers based on their amount of L2 immersion, which is regarded as a contributing factor

643   to define one's L2 developmental stage (see detailed discussion Fledge & Liu, 2001). Table

644   1 illustrates the distribution of the length of residence in English speaking countries, ranging

645   from 0 to 24 months (*M* = 4.6) with their age information (see Table 1). On the selection of

646   the audio, those 40 speakers' performance were chosen specifically because the speakers

647   answered to the similar prompt to the one used in the IELTS independent long-turn task,

648   which was also chosen for the prior IELTS pronunciation study by Isaacs et al. (2015).

649

Table 1
Length of residence and age for 40 Japanese speakers

Length of the stay (months)	*n*	Age	*n*
0	16	18–19	5
0–1	9	19–20	16
1–10	7	20–30	11
10–24	8	30–53	8
Total	40	Total	40

650

651   The IELTS speaking test takes an interview style with an examiner. In the long-tern

652   part, which is the second part of the test, the examiner provides a test-taker a prompt sheet

653   and asks the test-taker to speak after one or two minutes of preparation time. The prompt

654   used in the task is "*Describe the hardest and toughest challenge in your life*" and detail of

655   the prompt sheet is shown in Appendix C. Being in line with previous speech judgement

656   studies (e.g., Iwashita et al., 2008; Derwing & Munro, 1997), first 30 seconds of the entire

657   speech were excised from each of the 40 speaker's performance (approximately three

658   minutes each) for preparation of the ratings. Since this study's focus is on pronunciation, not

659   lexicogrammar or fluency, this study decided that 30 seconds is long enough to identify

660   pronunciation errors due to the empirical evidence such as Munro, Derwing, and Burgess's

661   (2010), which revealed that accented detecting ability of NEs was fairly reliable even with

662   single word. This is for avoiding raters' fatigue, which affect the accuracy of the rating result

663   due to its time-consuming nature, and completing all the error coding in the limited time

664   allocation.

665   **3.3.3 Error coding measures.** For the error coding of the 40 tracks of speech data,

666   eight pronunciation measures and six segmental measures were developed through two

667   studies (i.e., Trofimovich & Isaacs, 2012; Tutor, 2014). The eight phonological measures

668   contain segmental error, syllable error, misplacement of word stress, word stress absence,

669   misplacement of intonation, intonation absence. In order to include all the possible

670   phonological factors, which involve the IELTS pronunciation rating, the eight pronunciation

671   measures for current study were carefully selected from the research by Trofimovich and

672   Isaacs (2012) due to their in-depth coverage of phonological features. Regarding the six

673   problematic English segmental groups for NJs, the measures were adopted from Tutor's

674   (2014) priority list of problematic segmentals and suprasegmentals for NJs. This list was

675   particularly chosen because (a) the list offers grouped individual English segmentals, which

676   enable the current study to investigate the individual segmental errors in efficient way, (b)

677   the list represents the pronunciation problems unique to Japanese learner of English. All the

678   segment groups were derived from the contrastive analysis between Japanese and English as

679   well as the perception of native English and Japanese teachers of English, which enable the

680   current study to develop a educational guideline specific for Japanese learners of English,

681   and to examine universality of the results by comparing them with the prior linguistic

682   correlates study of IELTS pronunciation by Isaacs et al. (2015), whose L1 background of

683   speech samples were varied.

684   *3.3.3.1 Phonological measures.* Phonological measures are ranging from the

685   segmental level (individual vowels and consonants) to suprasegmental level (syllables,

686   words, and phrases). The number of errors in each measure was counted based on the

687   auditory evaluations by the researcher and two additional coders. The details of the measures

688   are as follows:

689

690   • Segmental errors (total): the errors caused by phonemic substitutions (e.g., *life*

691   articulated as *rife*).

692   • Syllable structure errors: the errors caused by vowel or consonant insertion (e.g., *it*

693   spoken as *ito*) and deletion (e.g., *year* articulated without the initial /y/).

694   • Word stress misplacement: the errors caused by the misplacement of primary

695   stress in polysyllabic words (e.g., *CHAL-lenge* spoken as *Chal-LEBGE*).

696   • Word stress absence: the errors caused by the absence of primary stress in

697   polysyllabic words (e.g., *CHAL-lenge* spoken as *chal-lenge*).

698   • Word stress error (total): the total errors of word stress misplacement and absence.

699   • Intonation misplacement: the errors in the adequate choice of pitch in the place

700   where certain type of tone is expected (e.g., falling tone at the end of statement

701   spoken in raising tone).

702   • Intonation absence: the errors of failing to produce any pitch movements (e.g.,

703   falling tone at the end of statement spoken with no tones).

704   • Intonation error (total): the total of errors in intonation misplacement and absence.

705

706   *3.3.3.2 Segmental measures.* Segmental measures comprise six groups of segmentals

707   (individual vowels and consonants) adapted from Tutor's (2014) problematic segmentals and

708   suprasegmental list (for the detailed segmental list see Appendix D). Through objective

709   analysis by the researcher and two additional coders, phonemes, which caused the segmental

710   errors, were identified and counted the numbers according to the group they belong to. The

711   details of the measures are as follows:

712

713   • Major segmentals (/ɹ, l, θ, v, ð/): the errors in articulating two approximants (/ɹ, l/),

714   and three fricatives (/θ, v, ð/) in words.

715   • Assimilation problems (/ʃi, si, ti/): the errors in articulating allophonic rules (/ʃi, si,

716   ti/) in words.

717   • Secondary segmentals (/f, æ, ʌ/): the errors in articulating a fricative (/f/), and two

718   vowels (/æ, ʌ/) in words.

719   • Diphthong problems (/aʊ, aɪ, oʊ, ɔɪ, eɪ/): the errors in articulating of diphthong

720   (/aʊ, aɪ, oʊ, ɔɪ, eɪ/) in words.

721   • Minor segmentals (/w, ŋ, h, n, p, t, k/, Contraction): the errors in articulating one

722   approximants (/w/), two nasals (/n, ŋ/), one fricatives (/h/), plosives (/p, t, k /),

723   allophonic rules (/ʃi, si, ti/) in words and contractions (e.g., won't, can't).

724   • Other segmentals: the errors in articulating the phonemes not included in the

725   group above.

726

727   **3.3.4 Speech rating.** Seven raters judged the 40 speeches of Japanese learners of

728   English based on the IELTS pronunciation rubric and compressibility. The rating sessions

729   were conducted individually in a quiet room (each rater participated in rating sessions on

730   different day and time) using a rating software *Praat* which shows two box-shaped 9-point

731   likert scales (1–9) for comprehensibility and the IELTS pronunciation judgement

732   respectively, raters then clicked the numbers to make the judgements (see Appendix E). The

733   raters listened to the audios through a pair of earphones connected to a personal laptop

734   computer.

735   On each rating, the raters were asked to listen to a whole audio (i.e., 30 seconds)

736   before they make the judgements on the two scales. The software played the speech only

737   once at randomised order, and the next speech was played right after the raters made the

738   judgements of the previous speech. All the ratings were recorded by clicking the numbers on

739   the two scales. Prior to the actual rating sessions, the language background questionnaire and

740   sufficient training were given to the raters. In the first phase, the researcher asked the raters

741   to complete the language background questionnaire shown in the computer screen in order to

742   obtain the information about the languages they speak and their degree of the familiarity

743   with Japanese and Japanese accented English. In the second stage, rater training session was

744   directed by the researcher to ensure that the raters fully adequately understood the IELTS

745   rubric and comprehensibility. In the training session, the physical copies of the IELTS

746   pronunciation rubric (9-point scale) and adequate time were provided with the raters in order

747   to help them fully understand the rubric. They were allowed to check the rubric anytime

748   during the rating process if necessary. The researcher then explained typical pronunciation

749   errors made by Japanese learners of English (e.g., syllabification problem caused by the

750   frequent vowel insertion) to raise the raters' awareness to the pronunciation errors. Then,

751   based on the definition used in the previous research of comprehensibility (e.g., Trofimovich

752   & Isaacs, 2012), the concept of comprehensibility (*how effortless to understand L2 speech*)

753   was introduced to the raters with a 9-point scale (1 = hard to understand, 9 = easy to

754   understand). For the comprehensibility rating, they were asked to make judgement

755   intuitively and use the scales flexibly. The raters listened to three sample audios and

756   complete practice judgement's with the rating software and discussed the result with the

757   researcher to confirm the raters' accurate understanding of the two scales.

758   **3.3.5 Error coding.** Firstly, eight pronunciation measures and six segmental

759   measures were coded by the researcher, who is a native speaker of Japanese and familiar

760   with Japanese-accented English due to the teaching experience to Japanese learners of

761   English. Subsequently, in order to ensure the correctness of the error coding by the

762   researcher and pursue a fine-grained objective measurement, the result of the first error

763   coding was re-coded by two Japanese coders. Those coders were given sufficient

764   instructions about the error detection before the actual codings (e.g., check the segmental

765   errors as if they were correcting their students' pronunciation based on their own ideal

766   pronunciation model). Two coders are experienced Japanese teachers of English (seven and

767   five years) and they are also knowledgeable about English phonology and phonetics with

768   master degrees in TESOL and Applied linguistics. For the coding, Japanese teachers'

769   subjective judgements were employed because native Japanese teachers of English are

770   assumed to be able to detect the influence of Japanese language easily in the L2 performance

771   compared to the native English teachers who do not speak Japanese as their L1. The results

772   of the additional coding did not show major difference from the first coding result,

773   suggesting that three coders' judgements were internally consistent.

774   **3.3.6 Data analysis procedure.** In line with the prior relevant literature**,** which

775   examined L2 speech from various perspectives (e.g., Tutor et al., 2015c), the quantitative

776   methods were carefully selected for the data analysis based on the research questions set in

777   this study. First, each speakers' mean IELTS pronunciation scores were calculated through

778   the averaging of seven raters' rating results in order to estimate each speaker's IELTS

779   pronunciation scores. For answering the first research question, forty mean scores were then

780   examined through Pearson's correlation to see the correlation with the coding results of eight

781   pronunciation measures and six problematic segmental groups. The second method, which

782   was applied to the second research question, involved in dividing of 40 rated speech samples

783   into three proficiency populations (i.e., beginner, intermediate, advanced). Then three groups

784   of confidence intervals were computed to identify the differences in possible ranges of

785   phonological and segmental error between three groups (Loewen & Plonsky, 2016). Finally,

786   in order to obtain the answers of the third research question, each speaker's

787   comprehensibility scores assigned by seven raters were averaged, and compared with the

788   average scores of the IELTS pronunciation via Pearson's correlation.

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

Chapter 4: Results

810   **4.1 Introduction**

811

812   This chapter presents an analysis of the data collected through the native judges'

813   rating session and the error coding result, and restates the research questions addressed in

814   Chapter 2. Based on the quantitative results obtained via the collected data from the rating

815   session and error coding, Pearson's correlation and confidence intervals are calculated to

816   identify the relationship between (a) the IELTS rating and eight pronunciation measures, (b)

817   the IELTS rating and six segmental measures, and (c) the IELTS rating and that of

818   comprehensibility. Each result is presented in response to the research questions.

819

820   **4.2 Rater Consistency**

821

822   An inter-rater reliability among seven NE raters was calculated through the averaging

823   of the values yielded through Pearson's correlation. The result supported a strong correlation

824   in their ratings ($r = .633$, $p > .1$), suggesting seven raters' judgements were consistent.

825

826   **4.3 Speech Data**

827

828   The rating results of 40 speech samples were manually classified into three

829   proficiency levels (i.e., beginner, intermediate, and advanced) in order to run a statistical

830   analysis for answering second research question (see Chapter 2). The result of the grouping

831   is presented in Table 2 below. As shown in the table, seven NJ speakers' mean scores were

832   Band 4 and another seven speakers' averaged scores were Band 5, fifteen of the speakers

833   had a mean score of Band 6, eight speakers had a mean score of Band 7, and three speakers

834   had a mean score of Band 8. Owing to the small sample sizes assigned to Bands 4, 5, 7, and

835   8 compared to that of 6, Bands 4 and 5 were conflated into beginner set ($n = 14$), Bands 7

836   and 8 were grouped together to make up advanced set ($n = 15$) whereas Band 6 remained as

837   it was and made up intermediate set by itself ($n = 11$).

838

839   Table 2

Results of Proficiency Level Grouping

Means of IELTS scores (Band score 0–9)	Number of speeches	Classification	Number of speeches in the proficiency groups
4	7	Beginner	14
5	7		
6	15	Intermediate	15
7	8	Advanced	11
8	3		

840

4.4 Correlation Between Pronunciation Features and the IELTS Pronunciation Rating

842

843   The first research question investigated the pronunciation features (i.e., segmental,

844   syllables, word stress, and intonation) of NJ learners of English, which were most crucial in

845   the IELTS pronunciation rating. In order to examine the strength of the association between

846   the mean rating results of 40 speech data and coded results of eight pronunciation features

847   (i.e., which features most affect the rating result), Pearson's correlations were computed.

848   Also, in order to obtain in-depth insight on segmentals for instructional purpose, the

849   correlation between the rating result and the error means of six problematic segmental

850   groups were examined through the correlation.

851   **4.4.1 Eight pronunciation measures.** As is presented in Table 3, the correlation

852   between eight pronunciation measures and the IELTS rating results were calculated. The

853   IELTS rating result showed significantly strong negative correlation with segmental errors (r

854   = -.590) and syllable errors (r = -.472). Similarly, the IELTS rating result was moderately

855   and negatively correlated with word stress absence (r = -.294) and total number of word

856   stress errors (r = -.304). The remaining four measures did not show any significance in the

857   correlation with the IELTS rating result: misplacement in word stress category) and

858   misplacement, absence, and total in intonation category. Especially, all the intonation

859   measurements showed positive correlation values compared to the rest of the measurements,

860   which all showed the negative correlation values.

861

862   Table 3
863   *Correlation Between Eight Pronunciation Measures and the IELTS Rating Results*

Pronunciation measures	r
Segmental errors	-.590[***]
Syllable errors	-.472[***]
Word stress misplacement	-.101
Word stress absence	-.294[*]
Word stress error total	-.304[*]
Intonation misplacement	.026
Intonation absence	.164
Intonation error total	.066

864   Note. $p* > .1, p** > .05, p*** > .01$, two-tailed

865

866   **4.4.2 Six segmental measures.** Pearson's correlation values of six problematic

867   segmental groups were computed (see Table 4). The IELTS rating result showed very strong

868   negative correlation with the major segmental group (r = -.473), and moderate negative

869   correlation with the minor segmental group (r = -.275). There was no significance in the

870   correlation between the IELTS rating and the rest of the segmental groups (i.e., assimilation

871   problems, secondary segmentals, diphthong problems, other segmentals).

36

872   In summary, these results suggest that pronunciation evaluation via the IELTS

873   pronunciation rubric considers segmental (especially major segmentals and secondary

874   segmentals) and syllable errors as crucial degrading factors. Yet not severely, word stress

875   errors, particularly the absence of the stress, impact on the negative evaluation of the

876   pronunciation scores. On the contrary, intonation errors do not lower the IELTS scores (i.e.,

877   raters seem to take a lax approach regarding any type of intonation errors). A summary of

878   the combined results of the correlation of all the measures with the IELTS rating according

879   to the significance is given in Appendix F.

880

881   Table 4
       Result of Pearson's Correlation between Six Segmental Groups and the IELTS
       Pronunciation Rating

Segmental error groups	r
1. Major segmentals	-.473[***]
2. Assimilation problems	-.121
3. Secondary segmentals	-.228
4. Diphthong problems	-.217
5. Minor segmentals	-.275[*]
6. Other segmentals	-.099

882   *Note*. $p^*>.1$, $p^{**}>.05$, $p^{***}>.01$, two-tailed

883

884   **4.5 Level Determinants of the IELTS Pronunciation Rating**

885

886   The second research question was engaged in identifying pronunciation aspects

887   (segmentals, syllables, word stress, and intonation), which most distinguish proficiency

888   levels of NJ learners of English in the IELTS pronunciation rating. In order to arrive at the

889   answer of this question, eight pronunciation features and six problematic English segmental

890   groups for NJ learners were examined according to the proficiency groups (beginner,

891   intermediate, advanced) determined by the IELTS pronunciation rating. Quantitative

892   approach was employed for the analysis of this research question (i.e., confidence intervals).

893   By using confidence intervals, the specified ranges of the errors each proficiency groups

894   make at 95% probability were compared.

895   For the calculation of confidence intervals of 95 %, all the rated speeches were sorted

896   into the three proficient groups according to the mean IELTS pronunciation scores they were

897   awarded (i.e., beginner for Bands 4 and 5, Intermediate for Band 6, advanced for Bands7 and

898   8). Then, on the basis of eight pronunciation measures and six segmental groups, mean error

899   and standard deviation of the coded error values (i.e., number of segmental, syllables, word

900   stress, intonation errors, and the number of segmental errors in group 1 to 6) of the three

901   proficiency groups were calculated to compute confidence intervals. The calculated means,

902   standard deviations, and confidence intervals of eight pronunciation measures and six

903   problematic segmental groups are presented in Table 5 and Table 6 respectively. Also, to

904   visually illustrate the relationship between three proficiency levels, the error distribution

905   figures are provided in the subsequent sections.

906   **4.5.1 Eight phonological measures.** Regarding the impact of segmental errors on

907   scoring distinction, significantly frequent error mean of 4.36 among beginners ($n = 14$, CI =

908   1.23, $SD = 2.34$), less frequent error mean of 1.29 among intermediate group ($n = 15$, CI =

909   0.73, $SD = 1.44$) and 1.67 among advanced group ($n= 11$, CI = 1.14, $SD = 1.94$) were

910   obtained. These results suggest the negative impact of segmental errors to distinguish

911   beginner from intermediate and advanced levels, while intermediate and advanced levels

912   were less distinguishable by segmental errors. A clear statistical significance of segmental

913   error effect on distinguishing beginner from the rest of the two proficiency levels could be

914   observed in confidence intervals at 95% probability in Figure 1: The lower limit of beginner

915   group [3.13–5.58] is not overlapped with the upper limit of intermediate [0.56–2.01] and that

916   of advanced level [0.52–2.81]. Therefore, segmental errors best distinguish beginner from

917   intermediate and advanced groups.

918

919 Table 5
920 *Summary of the Confidence Interval of Eight Pronunciation Measures vs. Three Proficiency*
921 *Level*

Pronunciation error measures	Proficiency level	n[a]	M	SD	CI	95% CI	
						LL	*UL*
Segmental	Beginner	14	4.36	2.34	1.23	3.13	5.58
	Intermediate	15	1.29	1.44	0.73	0.56	2.01
	Advanced	11	1.67	1.94	1.14	0.52	2.81
Syllable	Beginner	14	3.54	3.38	1.77	1.77	5.31
	Intermediate	15	1.62	1.66	0.84	0.78	2.46
	Advanced	11	0.82	1.25	0.74	0.08	1.56
Word stress (misplacement)	Beginner	14	0.92	1.19	0.62	0.30	1.55
	Intermediate	15	1.77	2.01	1.02	0.75	2.78
	Advanced	11	0.36	0.67	0.40	-0.03	0.76
Word stress (absence)	Beginner	14	2.57	2.17	1.14	1.43	3.71
	Intermediate	15	2.15	2.03	1.03	1.12	3.18
	Advanced	11	1.36	2.62	1.55	-0.18	2.91
Word stress (total)	Beginner	14	3.43	2.06	1.08	2.35	4.51
	Intermediate	15	3.87	2.64	1.34	2.53	5.20
	Advanced	11	1.73	0.67	0.40	1.33	2.13
Intonation (misplacement)	Beginner	14	0.23	0.60	0.31	-0.08	0.54
	Intermediate	15	0.46	0.78	0.39	0.07	0.85
	Advanced	11	0.27	0.65	0.38	-0.11	0.65
Intonation (absence)	Beginner	14	-	-	-	-	-
	Intermediate	15	-	-	-	-	-
	Advanced	11	0.09	0.30	0.18	-0.09	0.27
Intonation (total)	Beginner	14	0.23	0.60	0.31	-0.08	0.54
	Intermediate	15	0.46	0.78	0.39	0.07	0.85
	Advanced	11	0.36	0.67	0.40	-0.03	0.76

922 *Note*. CI = confidence interval; *LL*= lower limit; *UL*= upper limit.
923 n[a] = number of speech categorised as given proficiency group.
924

925   The contrasts of syllable effect on the rating showed distinctive result on the beginner

926 group with 3.54 mean error score ($n = 14$, CI = 1.77, $SD = 3.38$), while 1.62 error mean of

927 intermediate groups ($n = 15$, CI = 0.84, $SD = 1.66$) and 0.82 error mean of advanced group

928 ($n = 11$, CI = 0.74, $SD = 1.25$) are considered as relatively low. The results of 95 %

929 confidence intervals support this distinction of beginner level from intermediate and

930 advanced levels by presenting a relatively large range of errors of beginners [1.77–5.31] with

931 slight overlapping with intermediate group [0.78–2.46] and no overlapping with advanced

932 group [0.08–1.56] (see Figure 2). As for intermediate, its confidence intervals overlapped

933 with the lower limit of the beginner group's and the upper limit of the advanced group's

934 confidence intervals, suggesting that the syllable error does not clearly distinguish

935 intermediate and advanced but it does between beginners and advanced.

936
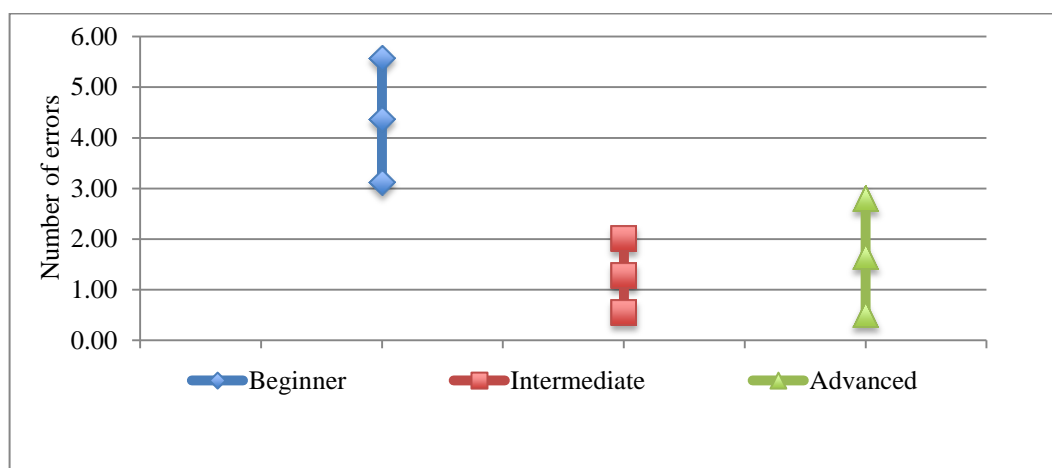
937
938 *Figure 1*. Confidence intervals of segmental errors.

939
940
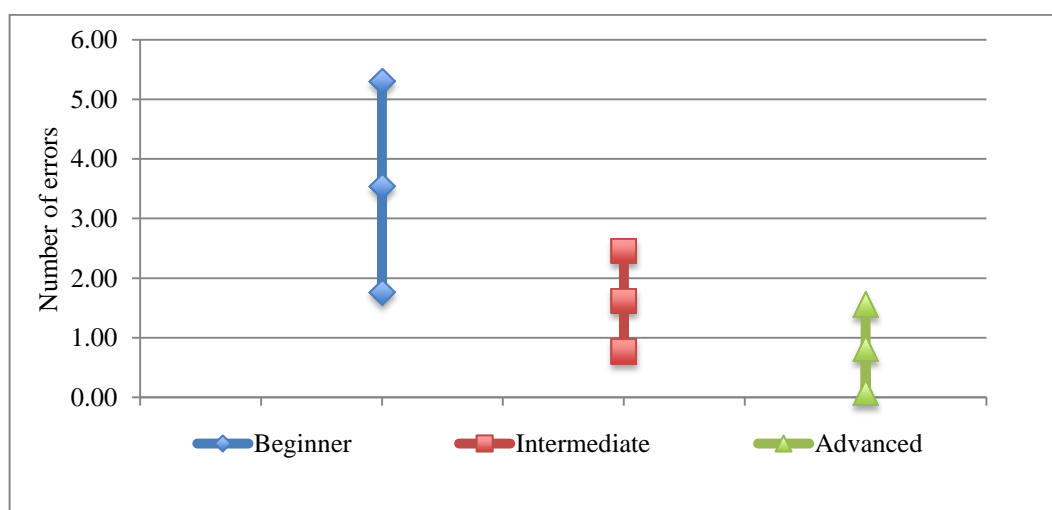
941
942 *Figure 2*. Confidence intervals of syllable errors.

943

944   In terms of word stress misplacement, relatively distinct difference can be seen

945   between highest error mean of 1.77 of intermediate ($n = 15$, CI $= 1.02$, $SD = 2.01$) and 0.36

946   of advanced levels ($n = 11$, CI $= 0.40$, $SD = 0.67$). Beginners ($n = 14$, CI $= 0.62$, $SD = 1.19$)

947   showed in-between error mean of 0.92, which shows no distinction between beginner and

948   intermediate levels. The result of confidence intervals (see Figure 3) yielded nearly

949   significant difference between intermediate level (0.75–2.78) and advanced level (-0.03–

950   0.76) with slight overlapping, while beginner group (0.30–1.55) is overlapping to a large

951   extent with intermediate, suggesting word stress misplacement weakly contributes to the

952   discrimination between intermediate and advanced groups.

953
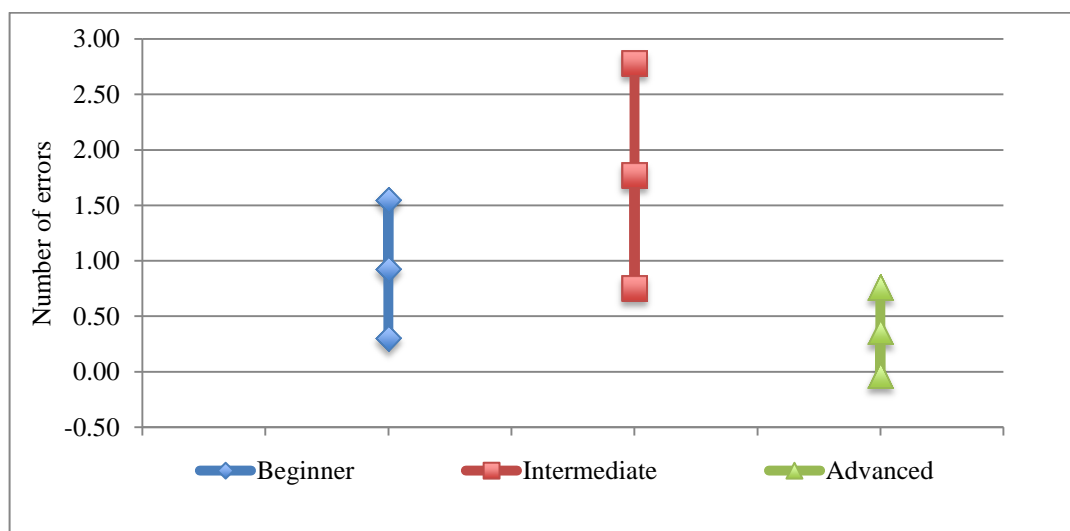
954
955   *Figure 3.* Confidence intervals of word stress misplacement.
956

957   With respect to the absence of the word stress, a gradual drop in the number of error

958   mean values was seen throughout the three proficiency levels, with 2.57 for beginner ($n =$

959   14, CI $= 1.14$, $SD = 2.17$), 2.15 for intermediate ($n = 15$, CI $= 1.03$, $SD = 2.03$) and 1.36 for

960   advanced groups ($n = 11$, CI $= 1.55$, $SD = 2.62$). However, analysis of 95 % confidence

961   intervals demonstrated that absence of the word stress has no significance as score

962   determinant between three proficiency levels, where all of the confidence intervals are

963 overlapping: beginner [1.43–3.71], intermediate [1.12–3.18], and advanced levels [-0.18–

964 2.91], respectively (see Figure 4). In summary, word stress absence does not act as a crucial

965 level determinant factor.

966
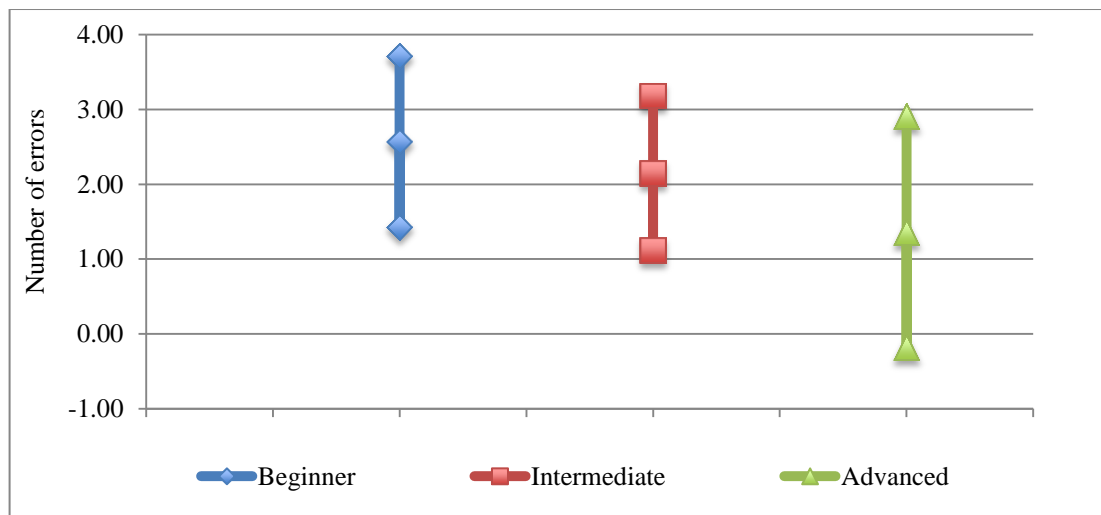
967
968 *Figure 4.* Confidence intervals of word stress absence
969

970   The overall word stress error (i.e., misplacement and absence of the stress in words)

971 revealed the mild contribution of word stress to distinguish the proficiency levels. The

972 calculated error means yielded 3.43 for beginner group ($n = 14$, CI = 1.08, $SD = 2.06$), 3.87

973 for intermediate group ($n = 15$, CI = 1.34, $SD = 2.64$), and 1.73 for advanced group ($n = 11$,

974 CI = 0.40, $SD = 0.67$), demonstrating clear distinction between intermediate and advanced

975 groups. This result was also supported by 95 % confidence intervals illustrated in Figure 5,

976 which show no overlapping between intermediate level [2.53–5.20] and advanced level

977 [1.33–2.13]. Major overlapping is seen between beginner level [2.35–4.51] and intermediate

978 levels [2.53–5.20], which indicate that beginner and intermediate levels are not

979 discriminated by word stress errors. Thus, word stress error in total is regarded as a strong

980 determinant to discriminate advance from intermediate groups but it is not the case between

981 intermediate and beginner groups.

982
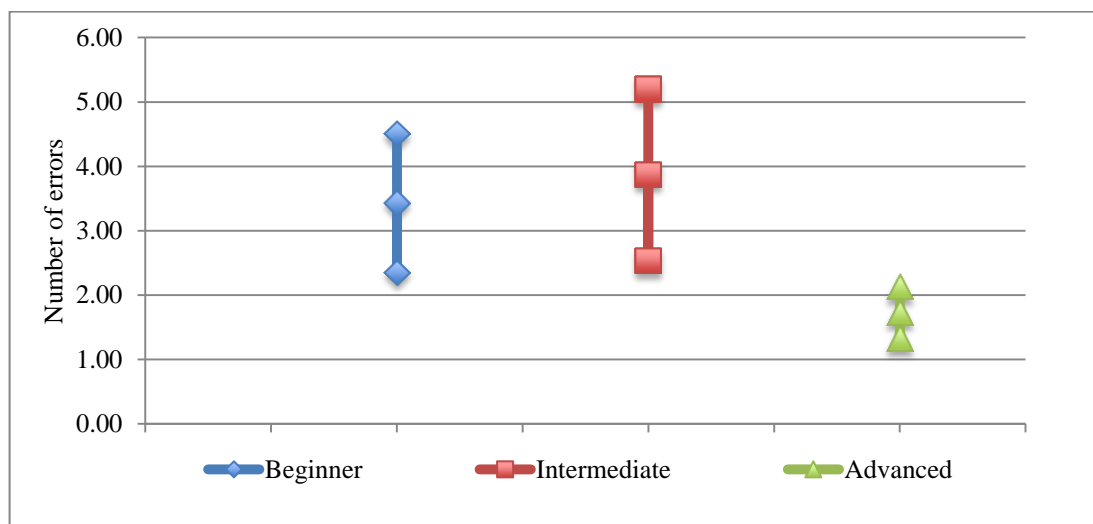
983
984 *Figure 5.* Confidence intervals of total errors of word stress

985

986   The means of intonation misplacement between the three proficiency groups did not

987 show any significant distinctions: 0.23 for beginner ($n = 14$, CI = 0.31, $SD = 0.60$), 0.46 for

988 intermediate ($n = 15$, CI = 0.39, $SD = 0.78$) and 0.27 advanced ($n = 11$, CI = 0.38, $SD =$

989 0.65). As is shown in Figure 6, 95 % confidence intervals of three proficiency groups are

990 overlapping each other and intermediate group [0.07–0.85] is higher than that of beginner

991 level [-0.08–0.54] and advanced level [-0.11–0.65]. These results suggest that intonation

992 misplacement error does not discriminate the differentiate proficiency levels on the IELTS

993 pronunciation rating.

994   Regarding the intonation absence errors, the error means were zero with beginner (n

995 $= 14$, CI = 0, $SD = 0$) and intermediate levels ($n = 15$, CI = 0, $SD = 0$), whereas 0.09 of

996 advanced level ($n = 11$, CI = 0.18, $SD = 0.30$) showed slightly higher mean. As is illustrated

997 in Figure 5, 95 % confidence intervals show zero range of beginner and intermediate groups

998 whilst advanced group [-0.09–0.27] shows rather wider range and overlap with both

999 beginner and intermediate groups. These results show that even the speeches of advanced

1000 group contained more intonation absence than that of beginner and intermediate, their

1001 speeches were rated higher. Thus, intonation absence does not impact on the IELTS
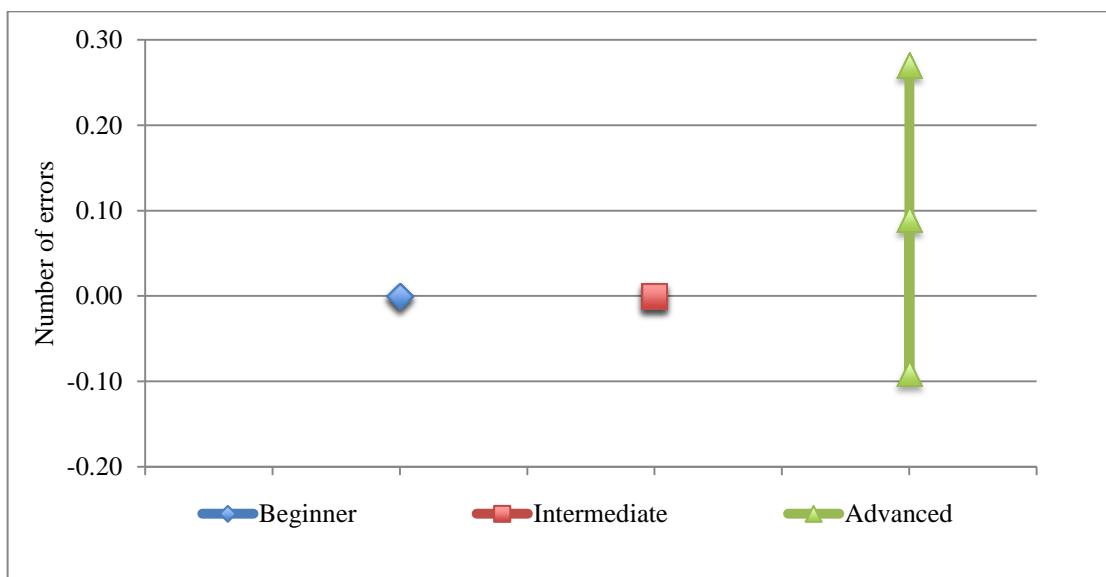
1002   pronunciation judgments.

1003
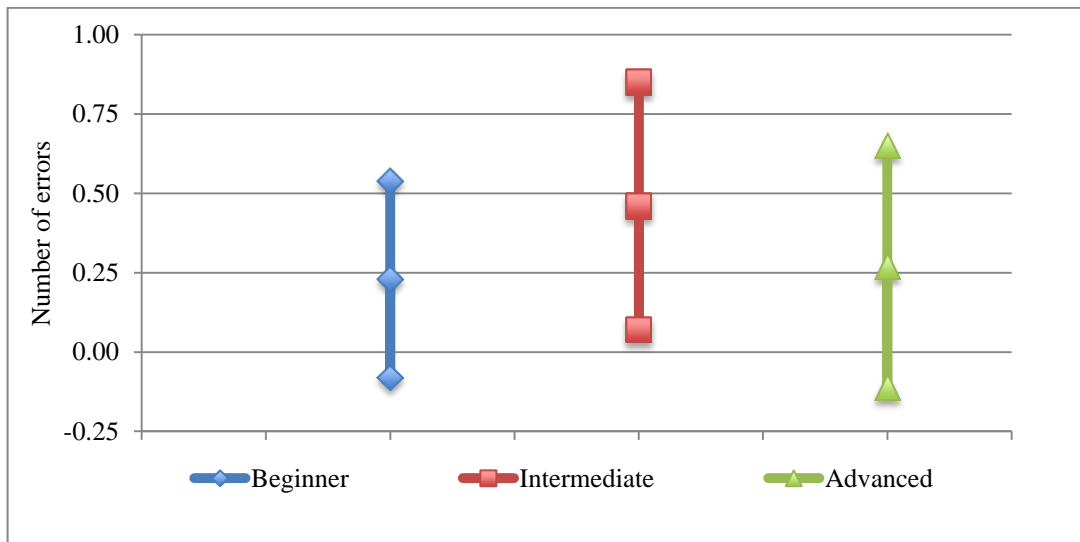
1004
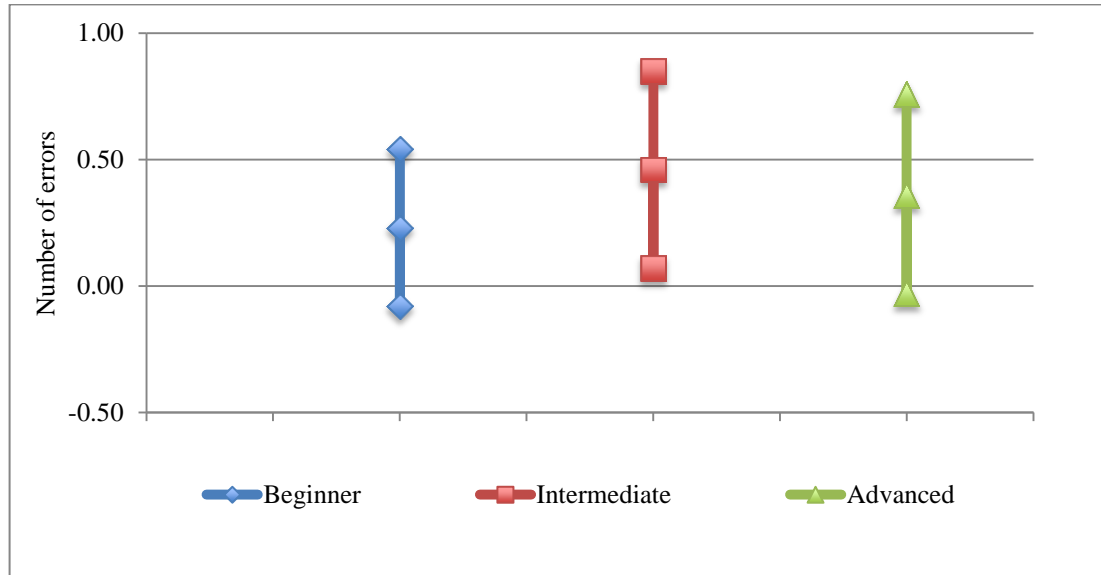1005   *Figure 6.* Confidence intervals of intonation misplacement errors

1006

1007
1008   *Figure 7.* Confidence intervals of intonation absence

1009

1010   The comparison of the error means of total intonation showed relatively high error

1011   mean of 0.46 with intermediate group ($n = 15$, CI = 0.39, $SD = 0.78$) compared to that of

1012   0.23 with beginner ($n = 14$, CI = 0.31, $SD = 0.60$) and 0.36 with advanced ($n = 11$, CI =

1013   0.40, $SD = 0.67$). With 95 % confidence intervals (see Figure 8), three proficiency groups'

1014   ranges are largely overlapped with each other: beginner [-0.08–0.54], intermediate [0.07–

1015   0.85], and advanced levels [-0.03–0.76]. This indicates that intonation error has no impact on

1016   discriminating three groups.

1017

1018
1019   *Figure 8.* Confidence intervals of total intonation errors
1020

1021   **4.5.2 Six segmental measures.** The results of confidence intervals at 95% level in

1022   six segmental groups are illustrated in Table 6. Regarding major segmentals, beginner group

1023   showed the highest error means of 2.57 ($n = 14$, CI = 0.84, $SD = 1.60$) compared to the other

1024   two groups: 1.00 of intermediate group ($n = 15$, CI = 0.51, $SD = 1.00$), and 1.27 of advanced

1025   group ($n = 11$, CI = 0.80, $SD = 1.35$). The comparisons of 95 % confidence intervals of three

1026   proficiency levels are illustrated in Figure 9. As is shown in Figure 9, the impact of errors of

1027   the first segmental groups proved significant in terms of discriminating beginners from

1028   intermediate level since there is no overlap between the beginner group [1.73–3.41] and that

1029   of intermediate [0.49–1.51]. On the contrary, the errors of the major segmentals do not

1030   distinguish intermediate and advanced because their confidence intervals are almost the

1031   same: intermediate level [0.49–1.51] and advanced level [0.48–2.07]. In summary, first

1032   group segmental is a strong determinant to separate beginner from intermediate and

1033   advanced groups.

1034 Table 6

1035 *Summary of Confidence Intervals of Six Segmental Measures*

						95% CI	
Segmental measures		n^a	*M*	*SD*	CI	*LL*	*UL*
Major segmentals	Beginner	14	2.57	1.60	0.84	1.73	3.41
	Intermediate	15	1.00	1.00	0.51	0.49	1.51
	Advanced	11	1.27	1.35	0.80	0.48	2.07
Assimilation problem	Beginner	14	0.21	0.43	0.22	-0.01	0.44
	Intermediate	15	0.20	0.41	0.21	-0.01	0.41
	Advanced	11	0.09	0.30	0.18	-0.09	0.27
Secondary segmentals	Beginner	14	0.14	0.53	0.28	-0.14	0.42
	Intermediate	15	-	-	-	-	-
	Advanced	11	-	-	-	-	-
Diphthong problems	Beginner	14	0.21	0.43	0.22	-0.01	0.44
	Intermediate	15	-	-	-	-	-
	Advanced	11	0.09	0.30	0.18	-0.09	0.27
Minor segmentals	Beginner	14	0.57	1.13	0.59	-0.02	1.16
	Intermediate	15	0.47	1.13	0.57	-0.10	1.04
	Advanced	11	0.09	0.30	0.18	-0.09	0.27
Other segmentals	Beginner	14	0.43	0.94	0.49	-0.06	0.92
	Intermediate	15	0.27	0.59	0.30	-0.03	0.57
	Advanced	11	0.09	0.30	0.18	-0.09	0.27

1036 *Note*. CI = confidence interval; *LL*= lower limit; *UL*= upper limit.
1037 n^a = number of speech categorised as given proficiency group.
1038
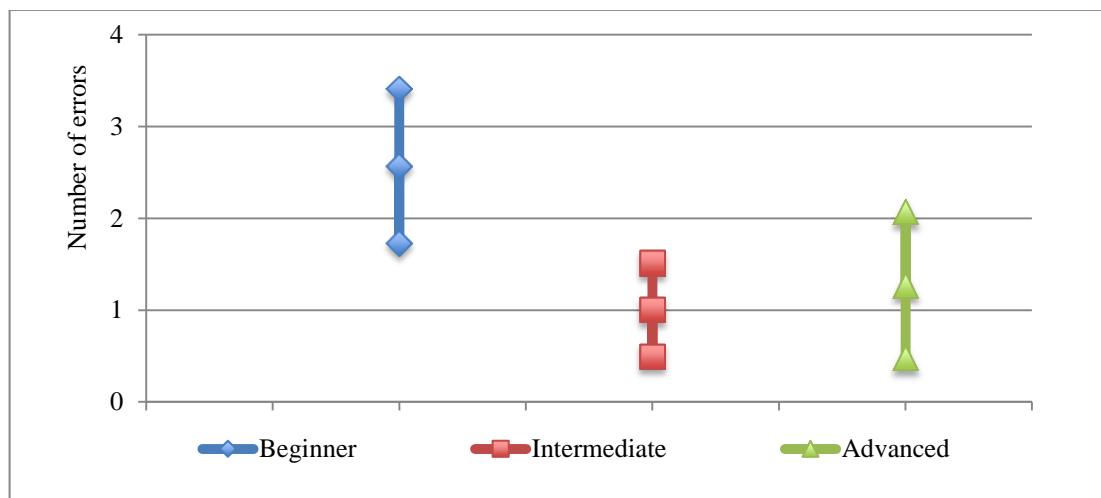
1039

1040 *Figure 9.* Confidence intervals of errors in major segmentals

1041

1042   With respect to the group comparison of assimilation problems, the error means of

1043   three proficiency groups revealed no significant difference between beginner and

1044   intermediate groups: 0.21 for beginners ($n = 14$, CI = 0.22, $SD = 0.43$), and 0.20 for

1045   intermediate group ($n = 15$, CI = 0.21, $SD = 0.41$), whereas slight difference is observed

1046   between 0.20 of intermediate and 0.09 of advanced groups ($n =11$, CI = 0.18, $SD = 0.30$).

1047   Similarly, 95 % confidence intervals yielded no significance between beginner and

1048   intermediate due to their major overlap: beginner level [-0.01–0.44], intermediate level [-

1049   0.01–0.41], while not major but some overlap is observed between intermediate level [-0.01–

1050   0.41] and advanced level [-0.09–0.27]. Thus, these results indicate that the assimilation

1051   problem group was not crucial factor to differentiate three groups.

1052
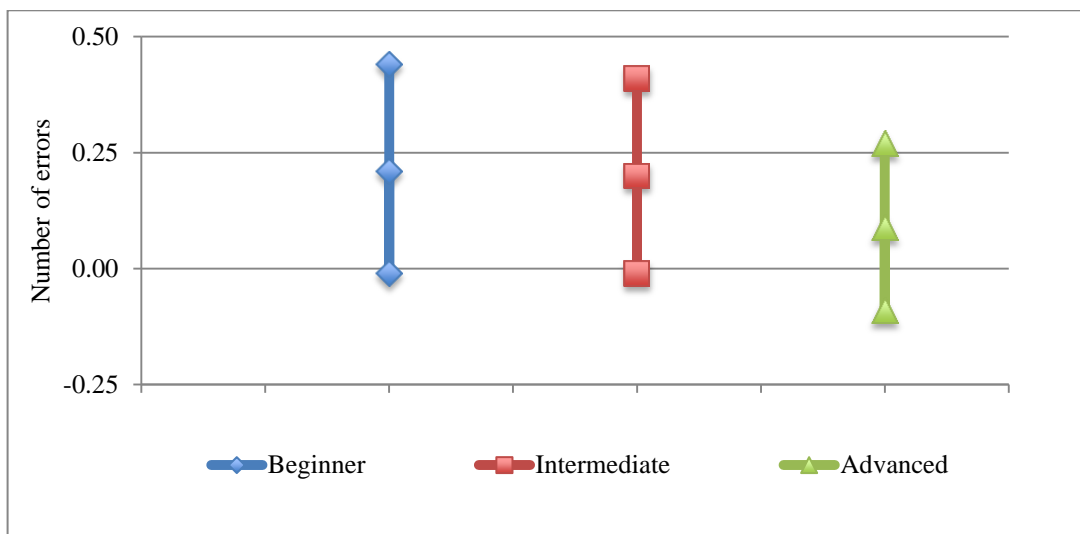
1053
1054   *Figure 10.* Confidence intervals of errors in assimilation problem
1055

1056   Considering the secondary segmental group, the error means of intermediate and

1057   advanced groups showed zero values compared to that of 0.14 for beginner ($n = 14$, *CI* =

1058   0.28, *SD* = 0.53). This indicates that beginner group is distinguishable from intermediate and

1059   advanced groups. As is illustrated in Figure 11, 95 % confidence intervals yield the complete

1060   overlap between intermediate and advanced, while beginner [-0.14–0.42] is overlapping with

1061 intermediate and advance groups at zero, indicating the distinction of beginner group from

1062 intermediate and advanced group is nearly significant. Overall, the secondary segmental

1063 group errors mildly impacted on the differentiation between beginner and intermediate

1064 groups but not between intermediate and advanced groups.

1065
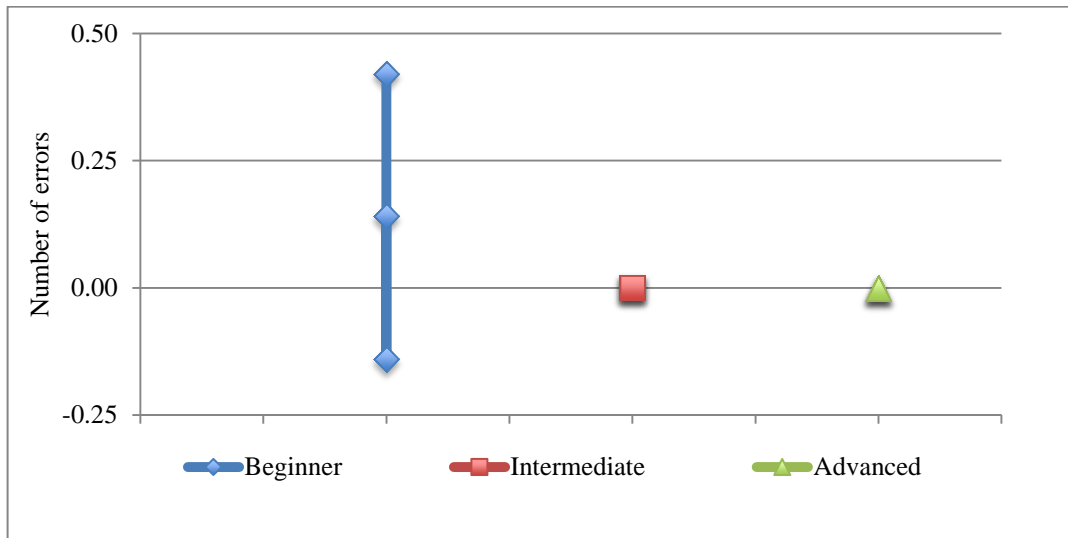
1066
1067 *Figure 11*. Confidence intervals of secondary segmentals
1068

1069 Diphthong problems showed nearly significant relationship between beginner and

1070 intermediate group comparison. As is illustrated in Table 6, the error mean of 0.21 for

1071 beginner group (*n* = 14, CI = 0.22, *SD* = 0.43) is higher than the zero value of intermediate

1072 group (*n* = 15, CI = 0, *SD* = 0). The error mean of 0.09 for advanced group (*n* = 11, CI =

1073 0.18, *SD* = 0.30) is slightly higher than intermediate group but lower than beginner group.

1074 Confidence intervals of 95 % (see Figure 12) revealed a partial overlapping between

1075 beginner [-0.01–0.44] and intermediate groups [0], which makes the distinction between

1076 beginner and intermediate nearly significant. However, advanced level [-0.09–027] and

1077 beginner level have a major overlap. Overall, these results suggest that diphthong problems

1078 might impact on discriminating beginner from intermediate groups but not on separation of

1079 intermediate from advanced groups.
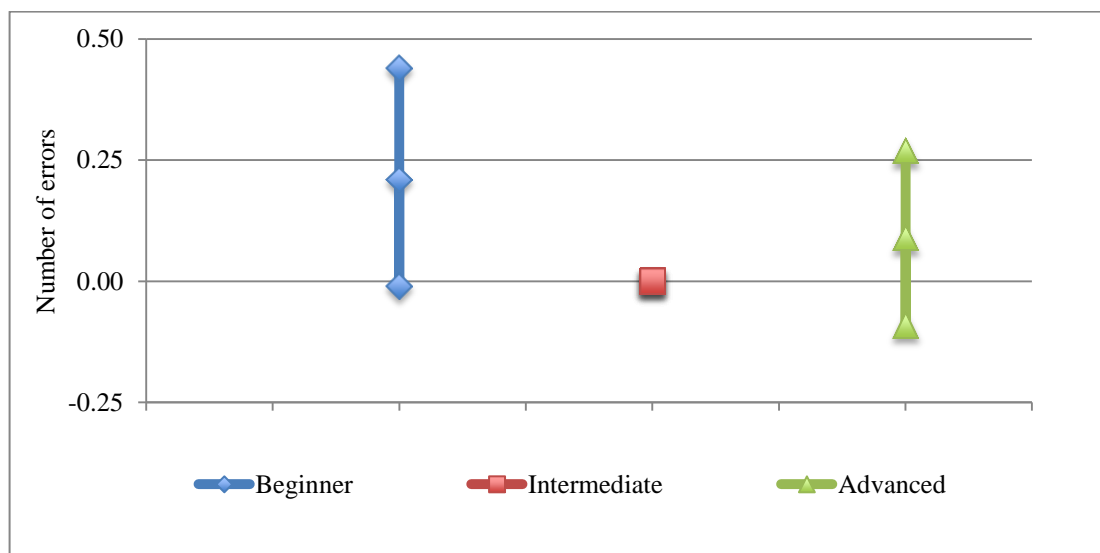
1081   *Figure 12.* Confidence intervals of errors in diphthong problems
1082

1083   Regarding the minor segmental group, the error means were 0.57 for beginner level

1084   ($n = 14$, CI = 0.59, $SD = 1.13$), 0.47 for intermediate level ($n = 15$, CI = 0.57, $SD = 1.13$),

1085   and 0.09 for advanced level ($n = 11$, CI = 0.18, $SD = 0.30$). These results shows relatively

1086   lower error mean of advanced group compared to that of beginner and intermediate groups.

1087   As is illustrated in Figure 13, 95 % confidence intervals revealed that no significance in the

1088   comparison between beginner group [-0.02–1.16] and intermediate group [-0.10–1.04] due

1089   to their major overlapping. However, the comparison of confidence intervals between

1090   intermediate and advanced groups [-0.09–0.27] almost reaches significance with slight

1091   overlapping. Therefore, although it is not significant, the error of minor segmental group

1092   appears to mildly contribute to the discrimination between intermediate and advanced group.

1093   Lastly, other segmentals showed gradual decline of the error mean from beginner to

1094   advanced groups: 0.43 for beginner ($n = 14$, CI = 0.49, $SD = 0.94$), 0.27 for intermediate (n

1095   = 15, CI = 0.30, $SD = 0.59$), and 0.09 for advanced levels ($n = 11$, CI = 0.18, $SD = 0.30$).

1096   Confidence intervals of 95 % of three proficiency groups are illustrated in Figure 14 and

1097   they show essential overlapping: beginner [-0.06–0.92], intermediate [-0.03–0.57], advanced

1098   levels [-0.09–0.27] respectively. Thus, no significance is observed in the impact of other

1099   segmental for level distinction.

1100
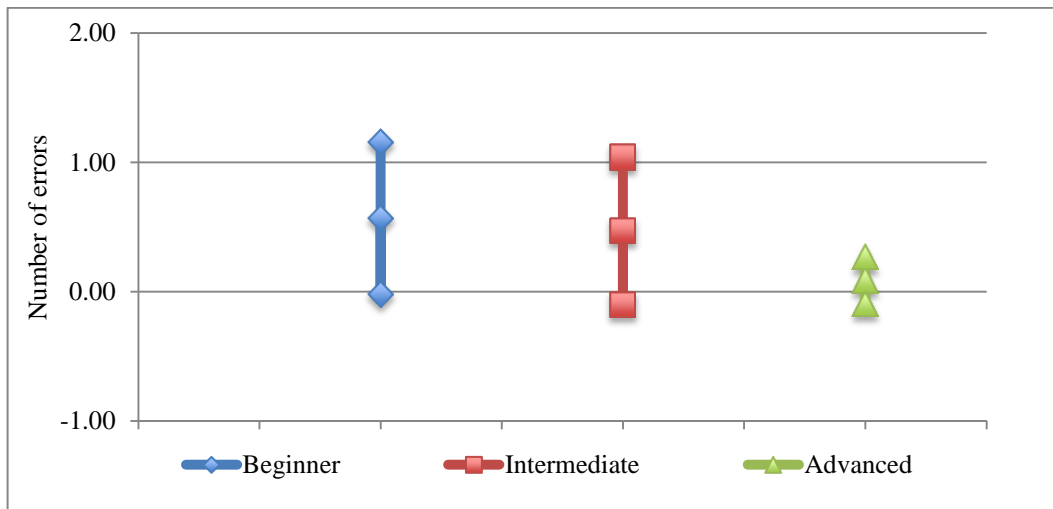
1101
1102   *Figure 13.* Confidence intervals of errors in minor segmentals
1103
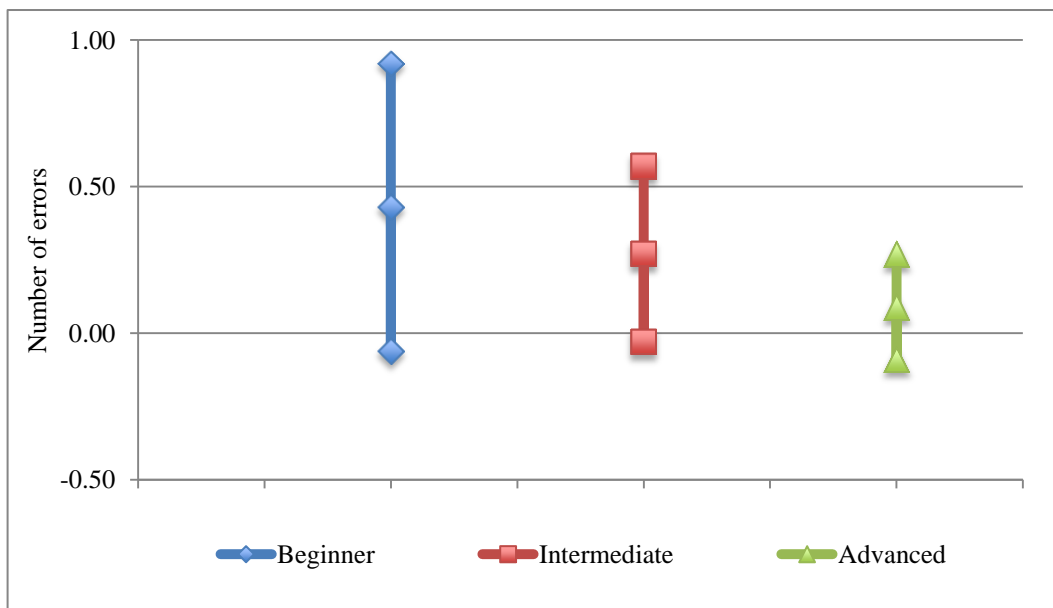
1104
1105   *Figure 14.* Confidence intervals of other segmentals
1106

1107   Taken together, beginner group is differentiated clearly by major segmentals and

1108   weakly by secondary segmentals and diphthongs while intermediate group is mildly

1109   discriminated from advanced group by minor segmentals. Nonetheless, other segmental

1110   groups are not strong determinant of separating three proficiency levels.

1111

4.6 Correlation between the IELTS pronunciation Rating and Comprehensibility

Judgement

    The third research question was set to identify the relationship between IELTS

pronunciation rating and that of comprehensibility rating. This question was answered

through a quantitative method, i.e., a statistical calculation. In order to see the correlation

between IELTS pronunciation rating result and that of comprehensibility, the Pearson's

correlation was used to examine the strength of the relationship between two variables ($-1 \leq r$

≥ 1). As a first step, mean IELTS score and mean comprehensibility score of each audio

were calculated. Then, the averaged rating results of each of the 40 recordings were inserted

to run person correlation coefficient. The correlation between the IELTS rating and

comprehensibility rating is yielded 0.958, suggesting significant positive correlation between

two rating patterns.
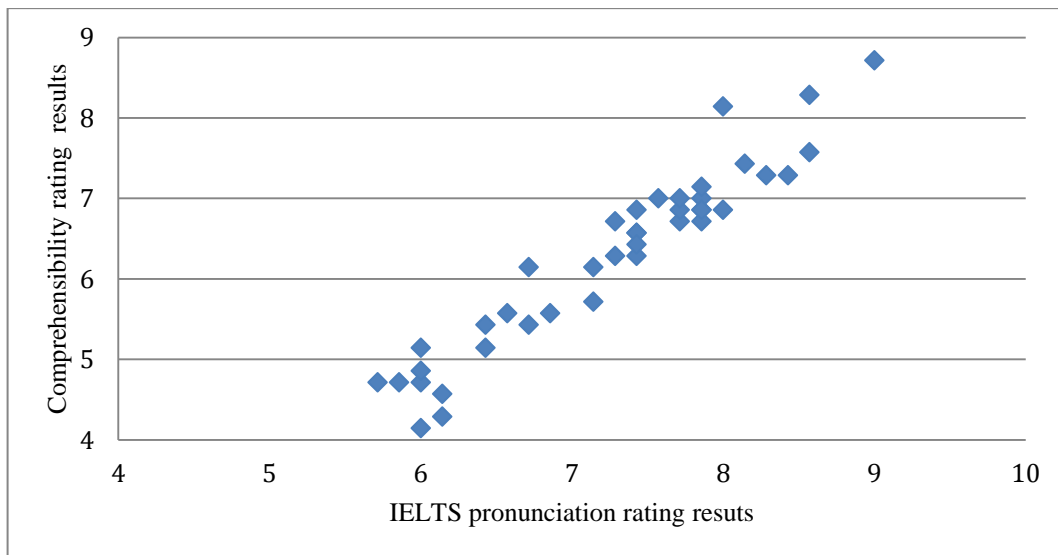
Figure 15. Result of Pearson's correlation between the IELTS pronunciation rating and
comprehensibility rating

    According to Evans and Over (1996), the value falls between 0.80 and 1.0, the

correlation between measured variable is regarded as very strong. Figure 15 illustrates

51

1132 positive correlation between the results of IELTS Pronunciation rating and

1133 Comprehensibility rating, confirming the strong link between two rating results.

1134

1135

1136

1137

1138

1139

1140

1141

1142

1143

1144

1145

1146

1147

1148

1149

1150

1151

1152

1153

1154

1155

Chapter 5: Discussion

5.1 Introduction

This chapter presents an extensive analysis of the crucial findings obtained through

the answers of the research questions. The analysis involves the discussions of the findings

in relation to the past relevant studies, which are followed by the brief re-statement of

research questions with the answers. The result of the first research question is discussed in

the first section with regard to the universal and L1 specific key features of IELTS

pronunciation rating as well as the uniqueness of IELTS pronunciation assessment from a

general L2 pronunciation syllabus. The second section emphasizes on the impact of L1 on

pronunciation error patterns, which result in the difference of pronunciation features to be

attended to by raters. The third section argues the importance of the further advancement in

the research in IELTS rating and pronunciation pedagogy.

5.2 First Research Question

The first research question considered the most crucial English pronunciation

features among segmentals and suprasegmentals (i.e., segmentals, syllable, word stress and

intonation) for Japanese IELTS candidates in the IELTS pronunciation assessment. This

question relates to the prioritisation within English pronunciation teaching and the

relationship between IELTS pronunciation measures and the two contrastive L2 speech

principles (i.e., comprehensibility and accentedness) regarding a specific L1 group (i.e.,

Japanese). A Pearson's correlation was computed to see the strength of the correlation

between errors of eight pronunciation measures and the IELTS pronunciation rating result of

1181 40 NJ speeches. The correlation analysis indicated that segmental errors severely affected

1182 the rating compared to that of suprasegmental errors. In fact, segmental errors and syllable

1183 errors showed the strongest/significant negative correlation with pronunciation rating.

1184 However, except for the mildly negative correlation between absence of word stress and

1185 rating result, the negative correlation of word stress misplacement, overall word stress errors,

1186 and all the intonation error variables (misplacement, absence and overall) on the rating were

1187 less significant. Regarding Pearson's correlation between the IELTS rating result and coded

1188 errors of six problematic segmental groups, a significantly strong negative correlation was

1189 found in errors of major segmentals. Furthermore, errors in minor segmentals showed a

1190 mildly strong correlation. The correlation values of other segmentals were not strong enough

1191 to reach significance. Taken together, the IELTS pronunciation rating was greatly affected

1192 by major segmental and syllable errors and mildly affected by word stress errors especially

1193 the absence of the expected stresses rather than misplacement. Other pronunciation features

1194 including assimilation problems, secondary segmentals, diphthong problems are regarded as

1195 minor factors but they still gave a negative effect on the rating. On the contrary, all the

1196 intonation errors showed positive correlations, indicating that Intonation errors were not

1197 considered as crucial factors in pronunciation ratings. Therefore, the following pedagogical

1198 suggestion for IELTS pronunciation teaching and learning can be made:

1199

1200   1. Firstly, teachers and learners need to focus on improving the production of major

1201   segmentals (/ɹ, l, v, θ, ð/) as well as syllabification to make sure not to insert

1202   unnecessary sounds to correct articulations or drop the sounds supposed to be

1203   articulated.

1204   2. Secondly, teachers and learners should concentrate on the practicing of minor

1205   segmentals (/w, ŋ, h, n, p, t, k/, Contractions) with particular emphasis on the

1206   placing of word stress when produce words and sentences.

1207

1208   3. As a final phase, secondary segmentals (/f, æ, ʌ/), diphthong problems (/aʊ, aɪ, oʊ,

1209   ɔɪ, eɪ/), assimilation problems (/ʃɪ, si, ti/) and other segmentals need to be

1210   introduced and practiced to further improve comprehensible articulation of

1211   English. Then the misplacements of word stress and the intonation errors (absence

1212   and misplacement) should be corrected.

1213

1214   The high correlation of segmental and syllable errors with IELTS rating in the

1215   current study showed a large disagreement with the findings in Isaacs et al. (2015), which

1216   lexicogrammatical factors showed highest correlation with the IELTS pronunciation rating.

1217   The cause of this difference can be due to two factors: (a) range of the linguistic

1218   measurements, (b) L1 effect. For the first factor, Isaacs et al. (2015) used a much wider

1219   range of measurement than the measurements in the current study. In fact, while the Isaacs et

1220   al. (2015) included grammatical accuracy, lexical richness speech chunking, and speech rate,

1221   the measurements employed for this study were limited to pronunciation features (i.e.,

1222   segmentals, syllables, word stress, and intonation). Therefore, the lack of lexicogrammatical

1223   aspect in the current study's measurements might have caused the high concentration of the

1224   rater's attention to the each of the pronunciation features such as segmental errors. However,

1225   a second factor to account for an anomaly in result is also possible as this study controlled

1226   the L1 background of the speech sample while this was not the case in Isaacs et al. (2015).

1227   Based on the evidence of the effect of learner's L1 on the L2 production being proven by the

1228   growing amount of literature in the area of L2 pronunciation development and

1229   comprehensibility (Tutor et al., 2015b), the result obtained in this research may have

1230   reflected the L2 pronunciation errors unique to Japanese learners of English, whereas the

1231   result of Isaacs et al. (2015), which contain the speech samples of multiple L1 backgrounds,

1232   seems to be regarded as more universal features of L2 pronunciation errors among learners

1233   from various L1 backgrounds.

1234   Another comparison of the result can be made with Tutor's (2014) propriety ranking

1235   of pronunciation features to predict the difference in the crucial features in IELTS

1236   pronunciation instruction and general English pronunciation instruction. This is because the

1237   ranking list in Tutor's (2014) was based on the NE and NJ teachers' perception of

1238   importance in their general English lessons to Japanese learners but not the IELTS test-

1239   taking classes. The findings in this study were partially consistent with the ranking presented

1240   in Tutor (2014) regarding the highest importance of major segmentals and syllabification

1241   problems (and cognates). This indicates the raters in this study and the teachers who

1242   participated in Tutor's (2014) study both perceived least comfort in understanding Japanese

1243   learners' speeches filled with major segmental and syllable errors. One of the possible

1244   accounts for the result of the weightiness of segmental and syllable errors is the difference in

1245   the phonetic systems of Japanese and English. For major segmentals (/ɹ, l, v, θ, ð/), four (/v,

1246   ð, l, ɹ/) out of five belong to English specific segmentals (/æ, f, v, q, ð, w, l, ɹ/). Therefore,

1247   these sounds would be the most difficult sounds for Japanese speakers to articulate

1248   accurately. In terms of syllable production, consonant clusters are found to be particularly

1249   difficult for Japanese learners. Due to the mora-timed nature of Japanese language, which

1250   place equal stress on each syllable (Ohata, 2004), Japanese speakers tend to insert

1251   unnecessary vowels after consonants in English words. In addition, due to the existence of

1252   numerous English loanwords, which are articulated in the Japanese sound system, Japanese

1253   learners are likely to conflate these loanwords with correct English pronunciations (see

1254   Riney & Anderson-Hsieh, 1993), indicating that use of loan words created incorrect

1255   syllabification in the audio samples examined in this study. The relatively crucial importance

1256   of major segmentals found in this study is also supported by the functional load theory

1257   (Brown, 1991; Catford, 1987; Kang & Moran, 2014; Munro & Derwing, 2006). Munro and

1258   Derwing (2006) revealed that segmentals of high functional load negatively and greatly

1259   impact on both L2 comprehensibility and accentedness. Since major segmentals include two

1260   high functional load segmentals (/ɹ, l /), the errors appeared to be fatal to L2 speech

1261   perception and might have induced the raters' negative judgments.

1262   On the contrary to the partial consistency with Tutor's (2014) study discussed above,

1263   the results of the relative importance of word stress, minor segmental and intonation for

1264   IELTS pronunciation rating did not fully support the priority ranking of Tutor (2014) based

1265   on teachers' perception. Word stress error was more crucial than assimilation problems in

1266   IELTS pronunciation rating whereas assimilation is regarded as more vital than word stress

1267   in Tutor's (2014) guideline. The relatively heavy weight absence of word stress on negative

1268   judgments might be derived from the raters' salient perception of monotonous sounds in the

1269   speech samples. Japanese learners' word articulations are due to the differences in the

1270   pronunciation system between English and Japanese. Compared to the vowel-focused

1271   English stress realization (i.e., longer and louder vowel of stressed syllable), stress is realized

1272   through higher pitch syllables (e.g., Gimson, 1989; Tsujimura, 2013; Vance, 1987). Minor

1273   segmentals (/w, ŋ, h, n, p, t, k /, Contraction) lead to lower rating consequence than

1274   secondary segmentals, diphthongs, and assimilation sounds in this study compared to their

1275   worst rank among the segmental groups in Tutor's (2014) implication. This result might be

1276   caused by the difference in perception between seven raters in this study and the NE/NJ

1277   teachers in Tutor's (2014) study: seven raters were more sensitive to the fricatives (/h/),

1278   plosives (/p, t, k/), nasals (/ŋ, n/) and approximal (/w/) errors. Intonation features appeared to

1279 be the lowest priority in the current study while they need to be taught before secondary

1280 segmentals, diphthongs and minor segmentals are covered in Tutor's (2014) implications.

1281 Seven native judges in the current study might be tolerant towards any types of intonation

1282 errors in their rating due to their own flexible use of intonations. As Levis (1999) pointed out

1283 in his investigation of NE natural speech sample in a corpus, native speakers seem not to

1284 always follow the intonation rules described in textbooks such as raising tones for "yes/no"

1285 questions. Whereas, NE/NJ teachers in Tutor's (2014) study might have strictly followed the

1286 intonation rules set in teaching materials and thus made intonation outrank secondary

1287 segmentals, diphthongs and minor segmentals. Overall, the results of the current study

1288 demonstrated different priority patterns from that of Tutor's (2014), suggesting that the

1289 raters perception based on IELTS pronunciation rubric and NE/NJ teachers' perceived

1290 pronunciation issues of native Japanese learners of English are different. Thus, when it

1291 comes to pronunciation instruction for IELTS, different approaches need to be taken it is

1292 suggested for teachers to adopt the priorities developed in the current study.

1293

1294 **5.3 Second Research Question**

1295

1296 The second research question examined English pronunciation aspects that most

1297 discriminate the proficiency levels of J learners in the IELTS pronunciation rating. This

1298 question is concerned with the salience of particular errors among three proficiency groups

1299 (beginner, intermediate, and advanced learners). Confidence interval of 95 % was employed

1300 for arriving at the answer to this research question. Forty rated speech samples were assigned

1301 to each proficiency group (beginner, intermediate, and advanced) according to the mean

1302 scores obtained through the averaging of the seven raters' rating results. Then, errors of the

1303 speeches in each group were calculated to yield confidence intervals, which are the ranges of

1304   the errors each proficiency group makes at 95% probability, were specified. Subsequently,

1305   these values were compared according to eight pronunciation features and six problematic

1306   segmental groups. The results of the group comparison regarding eight pronunciation

1307   features revealed that beginners were most distinguished by the segmental and syllable

1308   errors while other features did not show any significance in separating beginner from

1309   intermediate and advanced levels. Concerning intermediate learners, although individual

1310   word stress errors (i.e., absence and misplacement) were not significant enough to note, total

1311   errors of word stress impacted on discriminating them from advanced learners. In terms of

1312   the between-group comparison of six segmental groups, beginners were clearly

1313   discriminated by major segmental errors, and yet not significant, mildly separated by

1314   secondary segmental, and diphthong problem errors. In addition, however, it was not

1315   significant enough, minor segmentals weakly discriminated advanced from intermediate

1316   groups. Thus, the following IELTS pronunciation guideline for beginners and intermediate

1317   test-takers and teachers can be made:

1318

1319   1. In order to improve beginners to intermediate/advanced levels, major segmentals

1320   (/ɹ, l, v, θ, ð/) and syllables are the first features that need to be focused on among

1321   all the segmental groups. Then, secondary segmentals, and diphthongs should be

1322   instructed.

1323

1324   2. For the intermediate learners, word stress errors (misplacement, absence of the

1325   stress) need to be the propriety for the instruction. Subsequently, learners should

1326   work on the accurate production of minor segmentals.

1327

1328   Another noteworthy finding is, in terms of pronunciation features, the results were

1329   not fully consistent with the prior research of Isaacs et al. (2015) on IELTS pronunciation

1330   rating scales in separations of beginner, intermediate, and advanced. In their study, eight

1331   accredited IELTS examiners rated 80 candidates' speeches of various L1 backgrounds

1332   offered by Cambridge English. The ratings were conducted based on comprehensibility,

1333   segmental, prosodic, fluency, and lexicogrammatical measures and official examiner's

1334   version of the IELTS speaking rubric (confidential). Isaacs et al. (2015) found that Bands 5

1335   and 6 (beginner and intermediate in the current study) were best discriminated by speech rate

1336   and lexical richness, and the combined Bands 7 and 8 (equivalent to *advanced* in this current

1337   study) were differentiated from the combined Bands 5 and 6 by all the eight speech measures

1338   (i.e., comprehensibility, vowel and consonant errors, word stress, intonation, speech

1339   chunking, speech rate, lexical richness, and grammatical accuracy and sentence structure).

1340   Concerning the word stress errors, the current study and Isaacs et al. (2015) both found that

1341   word stress was responsible for distinguishing intermediate from advanced. Vowels and

1342   consonants (equivalent to segmentals and syllables), however, were major determinants for

1343   distinguishing intermediate from advanced groups (i.e., Band 6 from 7), while the present

1344   study revealed that they chiefly impacted on discriminating beginners from intermediate

1345   groups except for minor segmentals' contribution to the separation between intermediate and

1346   advanced groups. Furthermore, intonation errors bore no relation to the score discrimination

1347   in the current study albeit they appeared to be degrading factors for advanced group in Isaacs

1348   et al. (2015).

1349   The speculations on the little agreement of the findings with Isaacs et al.'s (2015)

1350   prior study described above involve two possible accounts. Firstly, the linguistic variables

1351   employed in the two studies were different. While their study considered the variables based

1352   on all the IELTS speaking criteria (fluency, lexis, grammar and pronunciation), the current

1353   study asked raters to only focus on pronunciation variables. The limited scope of the current

1354   study may have lead raters' attentions to the accuracy of sound articulations and resulted in

1355   finding the substantial contribution of segmentals for rating, whereas the raters of Isaacs et al.

1356   (2015), who reached judgments by considering various features of the speech and resulted in

1357   finding low importance of segmentals for the score judgement. Secondly, it is possible that

1358   the result of the pronunciation rating is attributed to the nature of L1 (i.e., Japanese). As a

1359   body of literature of L2 pronunciation acquisition has been proved that the L1 has a great

1360   influence on the quality of L2 production and one's comprehensibility (e.g., Tutor et al.,

1361   2015a; Tutor et al., 2015b). Thus, building on Isaacs et al.'s (2015) results, accuracy in word

1362   stress is likely to be a universally essential feature for advanced learners of various L1

1363   backgrounds (except for Farci speakers), while segmental features are more important for

1364   Japanese learners and Chinese speakers.

1365

1366   **5.4 Third Research Question**

1367

1368   The last research question dealt with the benefit of the IELTS pronunciation skill for

1369   attaining successful L2 communication. Since this study is in line with the researchers'

1370   belief that comprehensibility is the key facilitator of communication, comprehensibility

1371   rating was used to measure the indicator of success in L2 communication. Thus, this

1372   question investigated the correlation between IELTS pronunciation rating and the

1373   comprehensibility rating. The Pearson's correlation showed an extremely high correlation

1374   between IELTS pronunciation and comprehensibility rating, indicating that IELTS

1375   pronunciation skill is beneficial for L2 learners to improve their comprehensibility, which

1376   leads to success in L2 communication. This result does not support the findings of the prior

1377   study directed by Isaacs et al. (2015), which showed a weak correlation between IELTS

1378   rating and comprehensibility compared to other linguistics measures they used. The cause of

61

1379   the difference in the degree of correlation between the two types of judgement can be

1380   speculated in twofold: rater factor, and measurement factor. Firstly, the raters of the present

1381   study are different from that of Isaacs et al. (2015) in the way they were trained, which might

1382   cause a difference in judgement patterns. Isaacs et al.'s (2015) study concerned accredited

1383   IELTS examiners' judgement which reflected the criteria of the assessment rubric offered

1384   only to examiners, while the current study examined the judgement results produced by

1385   native speakers of English who are not official IELTS examiners and trained though the

1386   publicly available IELTS rubric, which is not as same as the one IELTS examiners consult.

1387   Secondly, as was discussed in the earlier section, the linguistic measurements used in Isaacs

1388   et al. (2015) and the present study are different. While Isaacs et al. (2015) employed

1389   linguistic measures spanning from pronunciation to lexicogrammar, the present study only

1390   used pronunciation measures. These differences in the measurement range might have

1391   affected the raters' degree of attentiveness to each linguistic measures and lead to differences

1392   in the relationship between IELTS pronunciation rating and comprehensibility.

1393   Despite the difference in the rubric and rating condition (using of only the

1394   phonological aspects), at least this study proved that measuring the construct of public

1395   IELTS pronunciation rubric is highly likely interpreted as a similar construct as

1396   comprehensibility by the judges in this study. This means that pronunciation instruction

1397   based on IELTS pronunciation rubric would help test-takers to improve their

1398   comprehensibility (i.e., skill for L2 communicative success). However, further research is

1399   certainly required with professional IELTS examiners to truly re-examine the relationship

1400   between degree of comprehensibility and IELTS pronunciation scoring.

1401

1402

1403

Chapter 6: Conclusion

6.1 Research Findings and Implications

The current study was designed to examine segmental and suprasegmetal correlates

and level determinants of the IELTS pronunciation rating. In addition, the study explored the

proximity of the rating to comprehensibility judgements to see the practicality of the

pronunciation skill for the IELTS speaking test.

The findings suggested that pronunciation errors, which were highly specific to

Native speakers of Japanese (NJs), were linked to the raters' judgements on the IELTS

pronunciation scale. A negative correlation with the ratings was saliently observed in the

production of English-specific segmentals (e.g., / ɹ, l, v, ð/) and syllabification by inserting

vowels after consonants, indicating that errors derived from first L1 influence (i.e., absence

of certain English sounds in Japanese phonetics, and over-application of Japanese

articulation of consonant and vowel combinations to English consonant clusters). These

features also distinguished beginner level from intermediate and advanced level of speakers.

Likewise, score-affecting errors caused by the difference in stress system between English

and Japanese (see Tsujimura, 2013) could also be seen in the word stress errors

(misplacement and absence) which showed a moderately negative correlation with

pronunciation rating. In effect, failing to place stress was a common error among all

proficiency levels, whilst errors in displacement of word stress were especially less frequent

among advanced speakers.

The results of the current study confirmed that learner's L1 background impacts on

the types of errors and rating results. This finding would inform L2 pronunciation instructors

about the importance of being aware of the L1 specific difficulties learners face not only for

1429   improving one's L2 comprehensibly, but also for improving IELTS pronunciation scores. In

1430   the case of the pronunciation syllabi developed in this study, they are certainly beneficial for

1431   Japanese IELTS test-takers.

1432

1433   **6.2 Research Limitations**

1434

1435   When considering the methodological respect of the current study, several limitations

1436   need to be addressed for future research. First, the scope to capture the pronunciation errors

1437   of the sample speech was limited to pronunciation features in the current study, while the

1438   study by Isaacs et al.'s (2015), which the current study builds on, has a much wider scope

1439   ranging from lexicogrammatical to pronunciation measures. This narrow scope prevented the

1440   study from fully capturing the complex relationship between test-takers' errors and IELTS

1441   pronunciation teaching, especially the possible influence caused by lexicogrammatical

1442   errors. Due to the relatively short length of each speech sample, this study was not able to

1443   collect sufficient amount of intonation errors to analyse its impact on the rating. Therefore,

1444   intonation errors need to be examined with much longer length of speech samples to

1445   adequately capture the all the types of intonation errors for analysis.

1446   Secondly, unlike Isaacs et al. (2015) and Tutor et al. (2015a) which have different L1

1447   backgrounds in their speech samples, the current study only focused on a specific L1

1448   background group. While the concentration on one specific language background enabled

1449   the study to generate L1 specific pronunciation syllabi, it limited the investigation into

1450   universally problematic segmentals and suprasegmentals, which can be measured though the

1451   comparisons of the result between different L1 background groups.

1452   Thirdly, the current study did not employ a qualitative approach to capture raters'

1453   thinking processes partially due to its sheer focus on the rating results, and partially due to

1454   the limited time allocation of the research. Thus, employing think-aloud protocol or a post-

1455   rating interview may have been a more accurate method in which to supplement the account

1456   of the results of rater judgements. Such qualitative approach is commonly employed in

1457   related studies as a *mix-method approach* (Creswell & Clark, 2007) where qualitative data is

1458   used to complement quantitative data (Isaacs et al., 2015).

1459   Lastly, the native judges hired in present study were not accredited IELTS judges but

1460   trained by the researcher based on publicly available IELTS rubric. Thus, admittedly, the

1461   rating results obtained in the present study would not be exactly the same as the results as

1462   professional IELTS examiners produced.

1463

1464

1465

1466

1467

1468

1469

1470

1471

1472

1473

1474

1475

1476

1477

1478

Reference List

1479

1480 Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker

1481   judgments of nonnative pronunciation and deviance in segmentais, prosody, and syllable

1482   structure. *Language Learning, 42*, 529-555.

1483 Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer [Computer software].

1484   Retrieved from http://www.fon.hum.uva.nl/praat/

1485 Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in

1486   Canada. *TESL Canada Journal, 19*, 51-61.

1487 Brinton, D. M. (2012). Pronunciation instruction. In A. Burns & C. J. Richerds (Eds.), *The*

1488   *Cambridge guide to pedagogy and practice in second language teaching*. New York:

1489   Cambridge University Press.

1490 British Council. (n.d.-a). IELTS assessment criteria: Understand your score. Retrieved from

1491   http://takeielts.britishcouncil.org/find-out-about-results/ielts-assessment-criteria/

1492 British Council. (n.d.-b). IELTS for Study. Retrieved from https://www.ielts.org/what-is-ielts/ielts-

1493   for-study/

1494 British Council. (n.d.-c). Who accepts IELTS? Retrieved from http://takeielts.britishcouncil.

1495   org/choose-ielts/who-accepts-ielts/

1496 Brown, A. (1991). *Teaching English pronunciation: A book of readings*. London: Routledge.

1497 Brown, A. (2006). An examination of the rating process in the revised IELTS speaking test. *IELTS*

1498   *Research Reports, 6*, 1-30.

1499 Callaway, D. R. (1980). Accent and the evaluation of ESL oral proficiency. In J. W. J. Oller & K.

1500   Perkins (Eds.), *Research in language testing* (pp. 102-115). New York: Newbury House.

1501 Catford, J. C. (1987). Phonetics and the teaching of pronunciation: A systemic description of

1502   English phonology. *Current perspectives on pronunciation: Practices anchored in theory*,

1503   87-100.

1504   Cobb, T. (n.d.). Compleat lexical tutor. Retrieved from

1505   Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research*.

1506   California: SAGE.

1507   Cunningham-Andersson, U., & Engstrand, O. (1989). Perceived strength and identity of foreign

1508   accent in Swedish. *Phonetica, 46*, 138-154.

1509   Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M., & Scott, J. H.

1510   (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front rounded

1511   vowels in L1 English–L2 French acquisition. *Second Language Research, 28*, 5-40.

1512   Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern Language*

1513   *Review, 59*, 547-567.

1514   Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in*

1515   *Second Language Acquisition, 19*, 1-16.

1516   Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A

1517   research-based approach. *TESOL Quarterly, 39*, 379-397.

1518   Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to

1519   communication. *Language teaching, 42*, 476-490.

1520   DeVelle, S. (2008). The 2007 revised IELTS pronunciation scale. *Cambridge ESOL Research*

1521   *Notes, 34*, 36-39.

1522   Eckman, F. R. (2004). From phonemic differences to constraint rankings: Research on second

1523   language phonology. *Studies in Second Language Acquisition, 26*, 513-549.

1524   Eckman, F., & Iverson, G. K. (2013). The role of native language phonology in the production of L2

1525   contrasts. *Studies in Second Language Acquisition, 35*, 67-92.

1526   Escudero, P., & Boersma, P. (2004). Bringing the gap L2 speech perception research and

1527   phonological theory. *Studies in Second Language Acquisition, 26*, 551-585.

1528   Evans, J. S. B., & Over, D. E. (2013). *Rationality and reasoning*. Hove, UK: Psychology Press.

1529 Feast, V. (2002). The impact of IELTS scores on performance at university. *International Education*

1530   *Journal*, 3, 70-85.

1531 Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*, 399-

1532   423.

1533 Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception.

1534   In N. O. Schiller & A. S. Meyer (Eds.), *Phonetics and phonology in language*

1535   *comprehension and production: Differences and similarities* (pp. 319-355). Berlin: Mouton

1536   de Gruyter.

1537 Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second

1538   language. *Studies in Second Language Acquisition*, *23*, 527-552.

1539 Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Effects of age of second-language learning on

1540   the production of English consonants. *Speech Communication, 16*, 1-26.

1541 Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language

1542   acquisition. *Journal of Memory and Language, 41*, 78-104.

1543 Gimson, A. C. (1989). *An introduction to the pronunciation of English*, London: Edward Arnold.

1544 Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of

1545   suprasegmentals. *TESOL Quarterly, 38*, 201-223.

1546 Harding, L. (2013). Pronunciation assessment. In C. A. Chapelle (Ed.), *The encyclopedia of applied*

1547   *linguistics* (pp.1-6). Oxford: Blackwell.

1548 Hayes, B., & Read, J. (2004). IELTS test preparation in New Zealand: Preparing students for the

1549   IELTS academic module. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in*

1550   *language testing: Research contexts and methods* (pp. 97-111). Mahwah, NJ: Lawrence

1551   Erlbaum.

1552 Hyatt, D., & Brooks, G. (2009). Investigating stakeholders' perceptions of IELTS as an entry

1553   requirement for higher education in the UK. *IELTS Research Reports, 10*, 17-68.

1554   Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in

1555   non-native English-speaking graduate students. *Canadian Modern Language Review, 64*,

1556   555-580.

1557   Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2

1558   pronunciation: Revisiting research conventions. *Language Assessment Quarterly, 10*, 135-

1559   159.

1560   Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility. *Studies in Second*

1561   *Language Acquisition, 34*, 475-505.

1562   Isaacs, T., Trofimovich, P., Yu, G., & Chereau, B. M. (2015). Examining the linguistic aspects of

1563   speech that most efficiently discriminate between upper levels of the revised IELTS

1564   pronunciation scale. *IELTS Research Reports Online Series*, *48*, 1-48.

1565   Issitt, S. (2008). Improving scores on the IELTS speaking test. *ELT Journal, 62*, 131-138.

1566   Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language

1567   speaking proficiency: How distinct? *Applied Linguistics, 29*, 24-49.

1568   Jenkins, J. (2000). *The phonology of English as an international language*. Oxford, UK: Oxford

1569   University.

1570   Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for

1571   English as an international language. *Applied Linguistics, 23*, 83-103.

1572   Kang, O. (2010). ESL learners' attitudes toward pronunciation instruction and varieties of English.

1573   In J. Levis & K. LeVelle (Eds.), *Paper presented at the Proceedings of the 1st*

1574   *Pronunciation in Second Language Learning and Teaching Conference* (pp. 105-118).

1575   Ames, IA: Iowa State University.

1576   Kang, O., & Moran, M. (2014). Functional loads of pronunciation features in nonnative speakers'

1577   oral assessment. *TESOL Quarterly, 48*, 176-187.

1578   Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and

1579   judgments of language learner proficiency in oral English. *The Modern Language Journal,*

1580   *94*, 554-566.

1581   Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Abor:

1582   University of Michigan.

1583   Lazaraton, A. (2005). Quantitative research methods. In E. Hinkel (Ed.), *Handbook of research in*

1584   *second language teaching and learning* (pp. 109-224). Mahwah, NJ: Erlbaum.

1585   Levis, J. M. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly, 33*, 37-63.

1586   Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL*

1587   *Quarterly, 39*, 369-377.

1588   Loewen, S., & Plonsky, L. (2016). *An A-Z of applied linguistics research method.* London:

1589   Palgrave.

1590   Major, R. C. (2008). Transfer in second language phonology. *Phonology and Second Language*

1591   *Acquisition, 36*, 63-94.

1592   Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition,*

1593   *21*, 81-108.

1594   Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the

1595   perception of native and foreign-accented speech. *Language and Speech, 38*, 289-306.

1596   Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the

1597   speech of second language learners. *Language Learning, 49*, 285-310.

1598   Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation

1599   instruction: An exploratory study. *System, 34*, 520-531.

1600   Munro, M. J., Derwing, T. M., & Burgess, C. S. (2010). Detection of nonnative speaker status from

1601   content-masked speech. *Speech Communication, 52*, 626-637.

1602   Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation

1603   teaching. *System, 42*, 258-269.

1604   Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American

1605   English vowels: influence of training sets. *Journal of Speech, Language, and Hearing*

1606   *Research*, 50, 1496–1509.

1607   Ohata, K. (2004). Phonological Differences between Japanese and English: Several potentially

1608   problematic. *Language Learning, 22*, 29-41.

1609   Oller, J. W., & Ziahosseiny, S. M. (1970). The contrastive analysis hypothesis and spelling errors.

1610   *Language Learning, 20*, 183-189.

1611   Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom.

1612   *TESOL Quarterly, 35*, 233-255.

1613   Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2:

1614   A review. *Journal of Phonetics, 29*, 191-215.

1615   Riney, T., & Anderson-Hsieh, J. (1993). Japanese pronunciation of English. *JALT Journal, 15*, 21-

1616   36.

1617   Riney, T. J., & Takagi, N. (1999). Global foreign accent and voice onset time among Japanese EFL

1618   speakers. *Language Learning*, *49*, 275-302.

1619   Rossiter, M. J. (2009a). Perceptions of L2 fluency by native and non-native speakers of English.

1620   *Canadian Modern Language Review, 65*, 395-412.

1621   Scales, J., Wennerstrom, A., Richard, D., & Wu, S. H. (2006). Language learners' perceptions of

1622   accent. *TESOL Quarterly, 40*, 715-738.

1623   Seedhouse, P., Harris, A., Naeb, R., & Üstünel, E. (2014). Relationship between speaking features

1624   and band descriptors: A mixed methods study. *IELTS Research Reports Online Series*, 30. 1-

1625   30.

1626   Sekiyama, K., & Tohkura, Y. (1993). Inter-language differences in the influence of visual cues in

1627   speech perception. *Journal of Phonetics.* 21, 427–44.

1628   Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants.

1629   *Language Learning, 41*, 177-204.

1630   Timmis, I. (2002). Native-speaker norms and international English: a classroom view. *ELT Journal,*

1631   *56*, 240-249.

1632   Tokumoto, M., & Shibata, M. (2011). Asian varieties of English: Attitudes towards pronunciation.

1633   *World Englishes, 30*, 392-408.

1634   Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism:*

1635   *Language and Cognition, 15*, 905-916.

1636   Tsujimura, N. (2013). *An introduction to Japanese linguistics.* Oxford: Blackwell.

1637   Vance, T. (1987). *An Introduction to Japanese phonology*. Albany, NY: State University of New

1638   York Press.

1639   Varonis, E. M., & Gass, S. (1982). The comprehensibility of non-native speech. *Studies in Second*

1640   *Language Acquisition, 4*, 114-136.

1641   Walker, R. (2010). *Teaching the pronunciation of English as a lingua franca*. Oxford: Oxford

1642   University.

1643   Wennerstrom, A. (2001). Intonation and evaluation in oral narratives. *Journal of Pragmatics, 33*,

1644   1183-1206.

1645   Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative

1646   contributions of F0 and duration. *Speech Communication, 55*, 486-507.

1647   Yates, L., Zielinski, B., & Pryor, E. (2011). The assessment of pronunciation and the new IELTS

1648   pronunciation scale. *IELTS Research Reports, 12*, 1-44.

1649   Young-Scholten, M. (1985). Interference Reconsidered: The Role of Similarity in Second Language

1650   Acquisition. *Selecta, 6*, 6-12.

1651   Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility.

1652   *System, 36*, 69-84.

1653

1654

1655

1656

1657

1658   4.

1659   **Appendix A**

1660   Publicly available IELTS pronunciation rubric (British Council, n.d.-a)

1661

IELTS™

SPEAKING: Band Descriptors (public version)

Band	Fluency and coherence	Lexical resource	Grammatical range and accuracy	Pronunciation
9	• speaks fluently with only rare repetition or self-correction; any hesitation is content-related rather than to find words or grammar • speaks coherently with fully appropriate cohesive features • develops topics fully and appropriately	• uses vocabulary with full flexibility and precision in all topics • uses idiomatic language naturally and accurately	• uses a full range of structures naturally and appropriately • produces consistently accurate structures apart from 'slips' characteristic of native speaker speech	• uses a full range of pronunciation features with precision and subtlety • sustains flexible use of features throughout • is effortless to understand
8	• speaks fluently with only occasional repetition or self-correction; hesitation is usually content-related and only rarely to search for language • develops topics coherently and appropriately	• uses a wide vocabulary resource readily and flexibly to convey precise meaning • uses less common and idiomatic vocabulary skilfully, with occasional inaccuracies • uses paraphrase effectively as required	• uses a wide range of structures flexibly • produces a majority of error-free sentences with only very occasional inappropriacies or basic/non-systematic errors	• uses a wide range of pronunciation features • sustains flexible use of features, with only occasional lapses • is easy to understand throughout; L1 accent has minimal effect on intelligibility
7	• speaks at length without noticeable effort or loss of coherence • may demonstrate language-related hesitation at times, or some repetition and/or self-correction • uses a range of connectives and discourse markers with some flexibility	• uses vocabulary resource flexibly to discuss a variety of topics • uses some less common and idiomatic vocabulary and shows some awareness of style and collocation, with some inappropriate choices • uses paraphrase effectively	• uses a range of complex structures with some flexibility • frequently produces error-free sentences, though some grammatical mistakes persist	• shows all the positive features of Band 6 and some, but not all, of the positive features of Band 8
6	• is willing to speak at length, though may lose coherence at times due to occasional repetition, self-correction or hesitation • uses a range of connectives and discourse markers but not always appropriately	• has a wide enough vocabulary to discuss topics at length and make meaning clear in spite of inappropriacies • generally paraphrases successfully	• uses a mix of simple and complex structures, but with limited flexibility • may make frequent mistakes with complex structures though these rarely cause comprehension problems	• uses a range of pronunciation features with mixed control • shows some effective use of features but this is not sustained • can generally be understood throughout, though mispronunciation of individual words or sounds reduces clarity at times
5	• usually maintains flow of speech but uses repetition, self correction and/or slow speech to keep going • may over-use certain connectives and discourse markers • produces simple speech fluently, but more complex communication causes fluency problems	• manages to talk about familiar and unfamiliar topics but uses vocabulary with limited flexibility • attempts to use paraphrase but with mixed success	• produces basic sentence forms with reasonable accuracy • uses a limited range of more complex structures, but these usually contain errors and may cause some comprehension problems	• shows all the positive features of Band 4 and some, but not all, of the positive features of Band 6
4	• cannot respond without noticeable pauses and may speak slowly, with frequent repetition and self-correction • links basic sentences but with repetitious use of simple connectives and some breakdowns in coherence	• is able to talk about familiar topics but can only convey basic meaning on unfamiliar topics and makes frequent errors in word choice • rarely attempts paraphrase	• produces basic sentence forms and some correct simple sentences but subordinate structures are rare • errors are frequent and may lead to misunderstanding	• uses a limited range of pronunciation features • attempts to control features but lapses are frequent • mispronunciations are frequent and cause some difficulty for the listener
3	• speaks with long pauses • has limited ability to link simple sentences • gives only simple responses and is frequently unable to convey basic message	• uses simple vocabulary to convey personal information • has insufficient vocabulary for less familiar topics	• attempts basic sentence forms but with limited success, or relies on apparently memorised utterances • makes numerous errors except in memorised expressions	• shows some of the features of Band 2 and some, but not all, of the positive features of Band 4
2	• pauses lengthily before most words • little communication possible	• only produces isolated words or memorised utterances	• cannot produce basic sentence forms	• Speech is often unintelligible
1	• no communication possible • no rateable language			
0	• does not attend			

Appendix B

1665   Language background questionnaire for raters (adopted from Tutor et al., 2016a)

1666
1667   Name _____
1668
1669   Age _____
1670
1671   Language Background
1672
1673   • Your birthplace _____
1674
1675   • Second language proficiency (e.g., beginner, intermediate, advanced, French, Spanish)
1676   1. _____
1677   2. _____
1678   3. _____
1679
1680   • Your parents' first language background
1681   (Father: _____, Mother: _____)
1682
1683   • Language spoken at home (generally)
1684   (English ____%; _____, ____ %)
1685
1686   • Have you ever had any hearing problems? (yes, no)
1687
1688   • Have you ever taken any linguistics classes (especially phonetics/phonology)? If yes,
1689   what kinds of classes?
1690   _____, _____, _____, _____,
1691   _____, _____, _____, _____,
1692
1693   • How long and where have you taught English?
1694   _____, _____, _____, _____,
1695   _____, _____, _____, _____,
1696
1697   • Have you ever visited Japan (yes, no)
1698   • If yes, how long did you stay? _____
1699   • Have you ever taken any Japanese course (yes, no)
1700   • If yes, how long did you study Japanese? _____
1701
1702   How much are you familiar with Japanese-accented English?
1703   (1   2   3   4   5   6   )
1704   Not at all   Neutral   Very much
1705
1706   How often do you have contact with native Japanese speakers?
1707   (1   2   3   4   5   6   )
1708   Very infrequent   Infrequent   Neutral   Often   Very often
1709

1710 **Appendix C**

1711 A prompt sheet used for speech elicitation for 40 speech samples in Tutor's (2011)

1712 unpublished corpus

1713

1714 Describe the hardest and toughest challenge in your life.

1715 Discussion points
- When? How old and where were you?
1716 - Why did you encounter this challenge?
- Why was it so challenging?
1717 - Did anybody (e.g., friends, parents) help you?

1718

1719

1720

1721

1722

1723

1724

1725

1726

1727

1728

1729

1730

1731

1732

1733

1734

Appendix D

Ranking list of problematic segmental groups for NJs adopted from Tutor (2014)

Segmental Measures	Segmentals	Priorities
1. Major segmentals	/ɹ/, /l/, /θ/, /v/, /ð/	↑
2. Assimilation problems	/ʃi/, /si/, /ti/	More
3. Secondary segmentals	/f/, /æ/, /ʌ/	important
4. Diphthong problems	/aʊ/, /aɪ/, /oʊ/, /ɔɪ/, /eɪ/	
5. Minor segmentals	/w/, Contraction (e.g., won't, can't), /ŋ/, /h/, /n/, /p/, /t/, /k/	Less important
6. Other segmentals	Other segmentals	↓

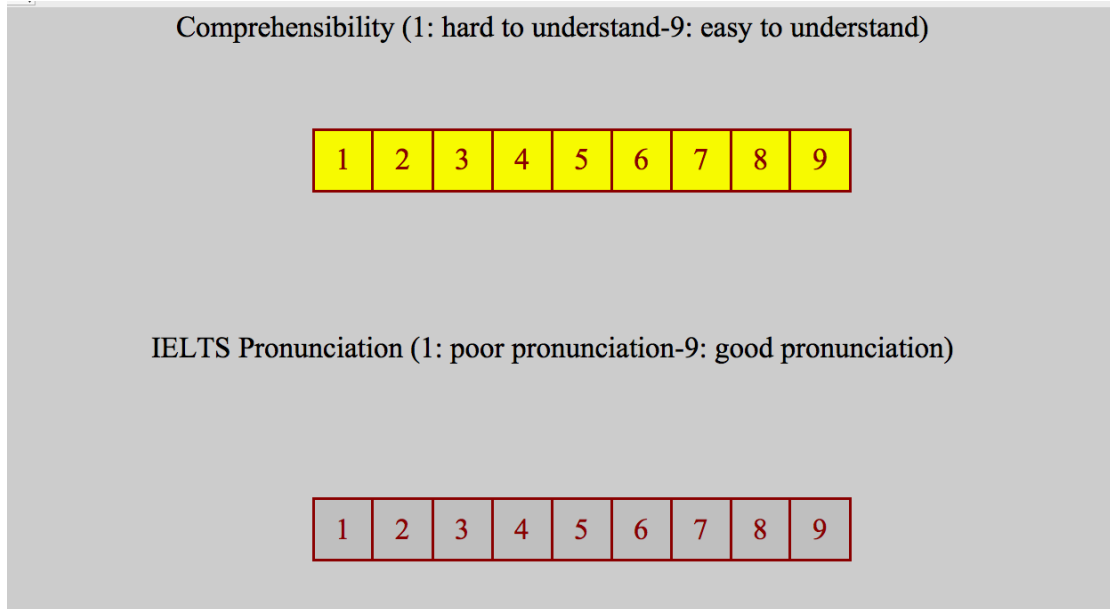*Note*. The measures are placed in order of importance (i.e., 1 is most and 9 is least important).

Appendix E

A screenshot of software Praat for rating

Comprehensibility (1: hard to understand-9: easy to understand)

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

IELTS Pronunciation (1: poor pronunciation-9: good pronunciation)

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

Appendix F

1775   Summary of Pearson's correlation between all the measures and the IELTS pronunciation

1776   rating

Phonological and segmental measures	r
Segmentals	-.590[***]
Major segmentals	-.473[***]
Syllables	-.472[***]
Word stress total	-.304[*]
Word stress absence	-.294[*]
Minor segmentals	-.275[*]
Secondary segmentals	-.228
Diphthong problems	-.217
Assimilation problems	-.121
Word stress misplacement	-.101
Other segmentals	-.099
Intonation absence	.164
Intonation total	.066
Intonation misplacement	.026

1777   *Note*. $p^* <.1$, $p^{**} <.05$, $p^{***}<.01$, two-tailed

1778

1779

1780

1781

1782

1783

1784

1785

1786

1787

1788

1789

1790