REVISITING THE PROBLEMATIC ENGLISH SOUNDS FOR PROSPECTIVE TURKISH EFL TEACHERS

Research article

(corresponding author)
İbrahim Halil Topal  
https://orcid.org/0000-0003-4220-3706
College of Foreign Languages, Gazi University, Ankara, Turkey
ibrahimtopal@gazi.edu.tr

İsmail Fırat Altay  
https://orcid.org/0000-0003-0567-1818
English Language Education, Hacettepe University, Ankara, Turkey
ifaltay@hacettepe.edu.tr

Biodata(s): No more than 40 words for each author.
İbrahim Halil Topal has been teaching English (A1-B2) for more than ten years at the College of Foreign Languages, Gazi University. His research interests include but not limited to educational technology, program evaluation, teacher education, and educational phonetics and phonology.

İsmail Fırat Altay an Assistant Professor at Department of ELT, Hacettepe University, Turkey. He offers lectures at undergraduate and graduate levels. Some of his interest fields are academic writing, pronunciation teaching, foreign language testing and non-verbal communication in language teaching.

Copyright © 2014 by International Online Journal of Education and Teaching (IOJET). ISSN: 2148-225X.

---

1 This paper is based on the doctoral dissertation titled ‘A Critical Evaluation of Pronunciation Teaching in an ELT Department in Turkey’ of the first author.
REVISITING THE PROBLEMATIC ENGLISH SOUNDS FOR PROSPECTIVE TURKISH EFL TEACHERS

İbrahim Halil Topal
ibrahimtopal@gazi.edu.tr
İsmail Fırat Altay
ifaltay@hacettepe.edu.tr

Abstract
Nonnative teachers of English as a foreign language (EFL), who must present role models, especially when their speech is the only source of input for their students (Richards, 2015), are a fifth group added to the list of Morley (1991), according to which teachers’ oral communication needs warrant a high level of intelligibility and thus requires them to receive special assistance with pronunciation (Celce-Murcia et al., 2010). Professionally speaking, EFL teachers must also possess certain teaching standards or qualifications concerning English pronunciation according to both national and international teaching frameworks (European Commission, 2011; Ministry of National Education, 2017; TESOL, 2019). Despite this essential requirement and the significance of pronunciation in oral communication (Pennington & Rogerson-Revell, 2019), pronunciation is deemphasized in teacher education settings (Baker, 2014). Given the professional and communicative salience of pronunciation, any pronunciation problems of EFL teacher candidates and practicing teachers should be surmounted, or related needs analyses should be performed. To this end, this study intended to re-examine the problematic segmental pronunciation features for Turkish EFL teacher trainees as to perception and production using a mixed-methods case study design (MM-CS). The findings revealed both similarities and differences with earlier studies. The study concluded that the teacher trainees were moderately accented, easily comprehensible, and comparatively intelligible.

Keywords: pronunciation; teacher trainees, segmentals; teacher education; needs analysis

1. Introduction
Given that correct pronunciation is key to effective communication (Pennington & Rogerson Revell, 2019), it is expected of language users to be well-resourced with the knowledge and use of intelligible pronunciation in verbal communication. This issue becomes far graver when language users in question are language teachers, since “they present role models which students may follow in their future use of the language and their practice as future teachers” (Council of Europe, 2001, p. 144). Both national and international teaching frameworks require the acquisition of certain skills and competencies concerning pronunciation and professionalism: Standards for English as a Second Language (ESL) / English as a Foreign Language (EFL) Teachers of Adults (TESOL, 2008), TESOL / National Council for Accreditation of Teacher Education (NCATE) P12 ESL Teacher Education Program (TESOL, 2010), The European Profiling Grid (European Commission, 2011), ACTFL Standards and CAEP Principles (ACTFL/CAEP, 2015), TESOL Pre-K-12 Teacher Preparation Programs (TESOL, 2019), and the General Competencies for Teaching Profession in Turkey (Ministry of National Education, 2017), to name a few. Since language teachers are considered role
models for language learners as aforementioned, they are expected to satisfy the requisites of language proficiency and professionalism (both theoretical and practical competencies about pronunciation) in accordance with the teaching standards and qualifications adopted nationally and internationally.

That being said, however, research in the related literature are at odds with what have been discussed thus far. In other words, pronunciation is underrepresented or overlooked in teacher education contexts owing to such various grounds as teachers’ knowledge, beliefs, attitudes, practices (Baker, 2014; Lindemann et al., 2016), revealing that there is trifling amount of pronunciation training (Sicola & Darcy, 2015) thus an urgency for training and professional development (Demirezen & Topal, 2015; Echelberger et al., 2016). Yet another reason for the de-emphasis of pronunciation in teacher education is its exclusion in teacher education curricula (Darcy, 2018; McGregor & Reed, 2018). Numerous scholars have suggested the inefficacy of teacher training programs for pronunciation instruction (Murphy, 2014; Baker & Burri, 2016). In the Turkish context, Yağız (2018) postulated that the curriculum failed to motivate teachers to teach pronunciation. The findings of a plethora of research carried out in the Turkish educational settings have lent support to the findings of both nationally- and internationally-conducted studies. More precisely, these research findings have demonstrated that both pre-service and practicing teachers experience articulation or other pronunciation-related problems (Bardakçı, 2015; Demirezen, 2016, 2017, 2020; Ercan, 2018 Yavuz & Keser, 2019), which might stem from particular reasons.

Considering the salience of pronunciation for effective and meaningful communication (Pennington & Rogerson-Revell, 2019), the stipulation of language proficiency (that includes a good command of proper pronunciation) across national and international teaching standards (European Commission, 2011; MoNE, 2017; TESOL, 2019), and the underrepresentation of pronunciation in teacher training curricula (Hişmanoğlu & Hişmanoğlu, 2013; Darcy, 2018; Munro & Derwing, 2019), this study aimed to discover the problematic English sounds (i.e., consonants and vowels) for Turkish EFL teacher trainees in terms of perception and production. It, therefore, intended to provide novel statistics about teacher trainees’ pronunciation competence and promote future pronunciation-associated studies concerned with the evaluation of teacher training curriculum within the Turkish context. Earlier academic work that focused on this subject is finite. Such Turkish scholars as Demirezen (2005a, 2005b), Hişmanoğlu and Hişmanoğlu (2011), and Ercan (2018) touched on the pronunciation problems of Turkish teacher trainees not merely to determine the problematic sounds but for their remediation. The present study, on the other hand, differs from the earlier ones in that it (i) considers the determination of problematic sounds relating to both perception and production, (ii) intends to provide novel data on the topic at hand, and (iii) bears particular methodological variances. In accordance with its objectives, three research questions were formulated in this study:

Problem of the Study

The primary and secondary problems of this study are as follows:

Problem: What are the problematic segmentals for Turkish teacher trainees as to perception, production, and remedial and expert judgments?
Sub-problems of the study

- What areas of English segmentals are potentially problematic for teacher trainees based on the remedial and expert judgment approaches?
- What English phonemes did the participants fail to perceive with regard to vowels and consonants?
- How well did the participants perform in terms of accentedness, comprehensibility, and intelligibility?

2. Literature Review

2.1. Accentedness, Comprehensibility, and Intelligibility

Three constructs come into play when it comes to identifying pronunciation errors: comprehensibility, accentedness, and intelligibility (Munro & Derwing, 1995). Accentedness refers to “a listener's perception of the ELL's speech as different from that of the listener's own language community” (Derwing & Munro, 2005, p.385). A native speaker might then ask “Is the speaker's pronunciation very different from the quality of pronunciation I am used to?” (Murphy, 2014, p. 261) while attending to a nonnative speech to identify whether or not it is accented. The second construct, comprehensibility, refers to ‘the listener’s perception of the degree of difficulty encountered when trying to understand an utterance’ (Munro et al., 2006, p.112). To find out whether a nonnative speech is comprehensible, a proficient listener might ask “Is it relatively easy or difficult for me to understand this speaker?” (Murphy, 2014, p.261). Munro and Derwing (1995) broadly define intelligibility as “the extent to which a speaker’s message is actually understood by a listener, but there is no universally accepted way of assessing it” (p.76). A proficient listener might ask “Do I understand the content of what this speaker has to say?” to determine the level of intelligibility of a nonnative speech (Murphy, 2014, p.261). Accentedness and comprehensibility are often assessed by listeners using a Likert-type scale, but intelligibility is often tested by having listeners transcribe what they hear word-for-word and computing an intelligibility score based on accurately transcribed words (Hansen Edwards et al., 2018). These three constructs, according to Murphy (2014), are technical terms referring to different aspects of pronunciation and reflect subjective perceptions of proficient listeners in relation to how accented, comprehensible or intelligible a nonnative speech is.

In this sense, it must be borne in mind that perceived intelligibility might be affected by such variables as “listener's L1, familiarity with nonnative English speech, receptivity, attentiveness, level of fatigue, familiarity with the topic being spoken about, etc.” (Murphy, 2014, pp. 258-259), which influence listeners' perceptions (Pickering, 2012). Whilst there is no clear indication as to which specific aspects of pronunciation are most crucial for intelligibility (Munro & Derwing, 1995; Darcy et al., 2012), explicit instruction of segmentals might assist learners with the development of a phonological awareness, which in turn might contribute to intelligibility (Saito, 2011b). For example, Collins and Mees (2013) selected seven aspects of pronunciation that they believed were the most significant for intelligibility and ranked them in the order in which they should be taught. In addition, according to Bent et al. (2007), intelligibility is substantially correlated with proper production of vowels and consonants in the English learners, in which eight segmentals were discovered to have a substantial impact on native speakers' speech perception.
Notwithstanding the lack of agreement on what aspects of pronunciation should be the goal in pronunciation instruction (Burri, 2016; Gao & Weinberger, 2018), foreign accent is an issue among non-native speakers, oft leading to communication breakdowns (Pennington & Rogerson-Revell, 2019). The role of accent or more specifically accurate pronunciation takes on a different role as far as language teachers are concerned in that they present role models for the language they teach (Council of Europe, 2001). They, therefore, are expected to have a good command of pronunciation as part of teaching qualifications as aforementioned.

2.2. Salience of Pronunciation for English Language Teachers

Language teachers are the primary source of information for many language learners within instructed language learning settings (Gallini & Barron, 2001). In certain contexts, they might even be the sole linguistic reference. For all learners, language teachers present role models in that they are considered the representatives of the language they teach (Council of Europe, 2001). In this sense, it would not be wrong to expect language teachers to be equipped well with the knowledge of and competence in the target language. In addition, the countries of the world mandate, to varying degrees, the acquisition of certain qualifications and teaching standards (e.g., European Commission, 2011; TESOL, 2019). Two concepts related to pronunciation in these teaching frameworks are language proficiency and professionalism, which indeed supports the previously stated fact that language teachers are required to be knowledgeable and competent in the target language pronunciation.

Pronunciation might also be deemed significant for language teachers considering the amount of training provided within the teacher education curricula. In the Turkish context, for instance, pronunciation is touched on merely within the scope of two undergraduate courses titled Listening and Pronunciation I/II as a part of the field education courses (Council of Higher Education, 2018). With a recent update in teacher education curriculum, the relevant boards of higher education institutions (i.e., departments of education) were assigned to determine the courses, curricula and credits in teaching programs (Council of Higher Education, 2020). This means that teacher education has been decentralized, which might raise serious concerns about the standardization and quality of education provided by 55 state and 16 private universities located in Turkey. For that reason, the state of pronunciation in language teachers should be followed and examined closely.

On the importance of pronunciation within teacher education curriculum, Hişmanoğlu and Hişmanoğlu (2011) expressed the following:

(i) Pronunciation instruction is important in foreign language teacher training. (ii) PI offered in the new curriculum is useful for raising the participants’ awareness of the importance of explicit pronunciation instruction. (iii) Raising competency and awareness in phonetics and phonology is the most prolific aspect of the course. (iv) Lack of divergent activities and technological applications are the most challenging sides of the course (Hişmanoğlu & Hişmanoğlu, 2011, p. 516).

Last but not least, pronunciation can be regarded significant for language teachers along with all other language learners, as it is required for effective communication (Levis & Wu, 2018; Nagle et al., 2018). The quote by Pennington & Rogerson-Revell (2019) underlines the salience of pronunciation for meaningful communication.
Pronunciation is not only a central and necessary aspect of communication to master, but in the best case is an aspect of spoken language that can result in positive interactions and add value and impact in aspects of life that depend on language and effective interaction with others. It is therefore an important basic as well as value-added factor for much of social, academic, and professional life centering on spoken language communication (Pennington & Rogerson-Revell, 2019, p.22).

Considering all this, the significance of pronunciation for all language learners – language teachers in particular – becomes apparent. The dearth of opportunities to receive pronunciation training within the teacher education curriculum in Turkey makes it valuable to conduct research on this topic within this context.

2.3. A Glimpse into the Turkish and English Sound Systems

Belonging to the Altaic family of the Turkic languages, Turkish is an agglutinative (Britannica, 2020) and a syllable-timed language while English belongs to the Indo-European family of the Germanic languages (Crystal & Potter, 2020) and is a stress-timed language. Both languages use Latin-script alphabet; however, they have orthographic and phonetic differences. Turkish has 29 letters (21 consonants and 8 vowels), seven of which have been modified to fit for the phonetic requirements of the language, whereas English has 26 letters (21 consonants and 5 vowels) but 44 sounds. The major differences between the two languages with regard to pronunciation is that words are pronounced the way they are written in Turkish, whilst this is not the case in English which has different letter-sound correspondences. Further details regarding the letters in both languages might be found in Tables 1-2 and Figures 1-2.

Table 1. Turkish consonants inventory

<table>
<thead>
<tr>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive and Affricate</td>
<td>p b</td>
<td>t d</td>
<td>tʃ dʒ</td>
<td>c ɟ</td>
<td>k g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f v</td>
<td>s z</td>
<td>ʃ ʒ</td>
<td>ɣ</td>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral Approximant</td>
<td>ɬ</td>
<td>ɮ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Adapted from Zimmer & Orgun, 1999, p. 154)
Table 2. American English consonants inventory

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p, b</td>
<td>t, d</td>
<td>k, g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>tʃ, dʒ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f, v</td>
<td>0, ð</td>
<td>s, z</td>
<td>ʃ, ʒ</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>l</td>
<td>j</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral Approximant</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Adapted from Ladefoged, 1999, p.41)

When Table 1 and 2 are examined, it can be observed that specific English sounds such as /θ/, /ð/ and /w/ do not exist in Turkish. It might, therefore, be claimed that such consonants might be problematic for Turkish speakers to pronounce.

As with the consonants, there are variations between English and Turkish in terms of vowels as well. Turkish vowels are categorized as front-back, round-unrounded, and high-low vowels while English vowels are divided into backness, height, roundedness, and tenseness-laxness. Such classification of the English vowels might pose problems for Turkish speakers in that the sound /ə/, for example, does not exist in the Turkish phonetic system. It, therefore, tends to cause trouble in pronunciation. The sound /æ/ is another variation in the phonetic systems of the two languages. The articulation and correct pronunciation of these and other potential problematic sounds might require extra effort and exertion by Turkish speakers of English. Apart from these individual sounds that might be tricky to pronounce, it should be borne in mind that there are sound changes that take place in English, such as assimilation and elision which might also necessitate careful attention by Turkish speakers – just like other nonnative speakers – for proper articulation and intelligible pronunciation.

(Adapted from Zimmer & Orgun, 1999)
Both English and Turkish have undergone phonological changes in their history. The Great Vowel Shift in English was a gradual change in the pronunciation of all long vowels wherever they occurred (Britannica, 2018), and the major/minor vowel harmony in Turkish can be defined "as a process in which the vowels in a word agree in relation to a certain phonetic or acoustic feature" (Hacıoğlu, 1994, p.289). Considering the effect of these changes in the phonology of the two languages and that “the sound system of the L1 exerts a great deal of influence on how second language (L2) pronunciation is learned” (Munro, 2018, p.267), it might be plausible to assert that Turkish speakers of English (teacher trainees in this case) might be expected to be aware of and knowledgeable about these phonological processes in line with the content knowledge that teacher trainees are expected to have (see MoNE, 2017, TESOL, 2019).

3. Method

3.1. Research Design

A MM-CS was utilized for the purpose of this study. Case studies "are a design inquiry found in many fields, especially evaluation, in which the researchers develops in-depth analysis of the case, often a program, event, activity, process or, one or more individuals. Cases are bounded by time and activity, and researchers collect detailed information using a variety of data collection procedures over a sustained period of time"(Creswell, 2015, p.14). A MM-CS, on the other hand, “is a type of mixed methods study in which the quantitative and qualitative data collection, results, and integration are used to provide in-depth evidence for a case(s) or develop cases for comparative analysis” (Creswell & Plano Clarke, 2017, p.116). Although case studies usually focus on individuals, they also can be some event or entity other than a single person. In this sense, the case for this study is English teacher trainees (first-year students) in a Turkish state and research university. For Ÿin (2018), case studies rely on multiple sources of evidence, with data needing to converge in a triangulation fashion. In line with this, the present study employed both qualitative and quantitative data (e.g., empirical findings, expert opinion, and perception and production tests). Before data collection, the Ethics committee approval was obtained for the research from Hacettepe University Ethics Committee dated 29.06.2021 numbered E-35853172-300-00001635768.

3.2. Context and Participants

Participants in this study were 49 undergraduate students (13 males, 36 females) enrolled in the English Language Teaching (ELT) program of a research university in Turkey. Their age ranged between 18 and 21. The participants were all Turkish, graduated from high school, and had been admitted to the department through the same steps; namely, two paper-based tests Temel Yeterlilik Testi (Basic Proficiency Test) and Yabancı Dil Testi (Foreign Language Test) and the addition of their high school cumulative general point average (CPGA). In addition, they were all freshmen at the ELT department. Taking these into account, they might be considered academically homogenous. The participants were chosen on a voluntary basis from the three sections of first-year class. They were informed about the research and their consent was gathered prior to data collection.
3.3. Instrumentation

3.3.1. Empirical Findings and Expert Judgment

Ellis (2006, p.30) suggested remedial and expert judgment approaches for determining the relative difficulty, learnability, and teachability of target linguistic elements for a specific group of L2 learners. The target feature in the remedial strategy can be chosen based on past empirical findings that show the feature is problematic for learners. In the expert judgment approach, researchers elicit the opinions of experienced L2 teachers to identify the problematicity of learners. Numerous benefits of seeking expert advice can be listed including being empirical, replicable, high face validity, and identification of psycholinguistically appropriate targets, as teachers typically consider learners' information-processing load (Saito, 2011a). Expert raters can be experienced or inexperienced English teachers (Saito et al., 2017); however, this study utilized the ratings of untrained listeners since ‘rating data from even untrained listeners reflect properties inherent in the stimuli and are therefore useful in the evaluation of speech samples’ (Derwing et al., 2004, p.672).

3.3.2. Perception of Spoken English (POSE) Test

The POSE test, designed by Shewell (2004), was used to identify the problem-causing sounds for the participants in terms of perception. It is described on its official website (https://posetest.com/) as a tool intended to assist nonnative English speakers (and their ESL/EFL teachers) in diagnosing speech perception, issues (Shewell, 2004). This test is validated and comprises five areas of focus: (i) vowels, (ii) consonants, (iii) word stress, (iv) intonation (sentence-final), and (v) sentence stress. However, only vowels and consonants were used for the purpose of this research. There are 34 vowel items and 84 consonant items in the test. Items of each category are supplied in minimal pairs. After hearing the word, test-takers must choose the correct word in minimal pairs. Those with a score of 75% or more are judged to have no issues with the relevant contrast (minimal pair). Because this is a computerized speech perception examination, it was given to the participants online. The participants were given some practice with the words in the test prior to the administration of the test to acquaint them with the unfamiliar vocabulary.

3.3.3. Productive Pronunciation Test

The segmental aspects of phonology can be tested in various methods, ranging from lexical-sentential item repetition to reading aloud (Hewings, 2004). For example, Elliot (1995) assessed participants' pronunciation accuracy by subjecting them to subject accuracy imitating speech at the discrete word level, subject accuracy mimicking pronunciation at the sentence level, and subject accuracy of written word pronunciation, and subject accuracy of pronunciation in spontaneous communication. Because there is no standardization in measuring the productive part of pronunciation, this study employed a productive pronunciation test (i.e., an elicitation paragraph). The participants were requested to record their voices while reading aloud a paragraph in a stress- and sound-reduced environment. They were given some time before taking the test to avoid “processing overload during the recording, which diverts their focus away from pronunciation and leads to dependence on their L1 habits” (Walker, 2005, p.553). A brief text named The Boy Who Cried Wolf—an altered version of one of Aesop's tales, which occurred in the Handbook of the IPA (International Phonetic Association, 1999, p.39) and suggested by Deterding (2006) were utilized, because it includes all the phonemic contrasts of the English language.
3.4. Procedures for Data Collection and Analysis

The data for the first research question were collected through remedial and expert judgment approaches (Ellis, 2006), by which analysis of empirical findings were performed and opinions of experts were collected on the problematic English segmentals via an online survey question asking them to rate problematic English segmentals from 1 (not problematic at all) to 9 (extremely problematic) for prospective English teachers. For the second research question, the data were acquired through the POSE test (Shewell, 2004) that assessed the participants’ success levels in perceiving English vowels and consonants. The data for the last research question were gathered via the productive pronunciation test that aimed to assess the participants’ proficiency levels in producing English sounds. The mean scores were calculated for the data corresponding to the first two research questions via Microsoft Excel. The data for the last research question were rated by seven raters using a nine-item Likert-scale (Derwing & Munro, 2005; Hansen Edwards et al., 2018) with regard to accentedness and comprehensibility and a transcription task for intelligibility (Derwing & Munro, 2005; Hansen Edwards et al., 2018). Seven raters with at least 10 years of teaching experience evaluated the speech samples of the participants. The ratings were put to intra-class reliability prior to analysis. A high degree of reliability (DeVellis, 2016) was found between the ratings for accentedness (α=.99 with a 95% confidence interval from .995 to .998, (F=48, 1200) =278,364, p=.000) and comprehensibility (α=.98 with a 95% confidence interval from .976 to 989, (F=48, 1200) =60,115, p=.000).

4. Findings and Discussion

RQ (1): What areas of English segmentals are potentially problematic for teacher trainees based on remedial and expert judgment approaches?

Remedial and expert judgment approaches (Ellis, 2006) were employed to answer the first research question. According to remedial approach, the target problematic sounds can be determined ‘based on previous empirical findings that have demonstrated the feature is problematic to learners’ (Ellis, 2006:30). In this regard, earlier empirical findings revealed that the following English consonants and vowels were problematic for Turkish EFL teacher candidates: /dʒ/, /ʃ/, /ʃ/, /ʒ/, /v/, /w/, /ŋ/, /r/, /ə/, /æ/, /ɑ/ (Demirezen, 2005a, 2005b, 2007, 2010, 2017); /θ/, /ð/, /ʊ/, /uː/, /eɪ/, /ε/, /i/, /ɪ/, /ʌ/, /ɔː/, /oʊ/, /æ/ (Hişmanoğlu, 2009; Hişmanoğlu & Hişmanoğlu, 2011), dark /l/ (Kahraman, 2013), /ɛ/, /θ/, /ŋ/, and /æ/ (Bardakçı, 2015); /θ/, /ð/, /w/, /v/, /ŋ/, /ɪ/, /ʊ/, and /ɔː/ (Ercan, 2018); /θ/, /ð/, /w/, /dʒ/ (Mahzoun and Han (2019).

The second approach (i.e., expert judgment), on the other hand, utilizes experienced L2 teachers’ opinions about the problematic structures (Saito, 2011a). In this study, the problematicity was the problematic English segmentals for prospective Turkish EFL teachers. To this end, 11 non-native experienced L2 instructors (6 females, 5 males), five of whom taught pronunciation course before, with 11-30 years of teaching experience were consulted and provided with a list of English sounds, among which they were asked to rate each with a score between 1-9 (1= not problematic at all, 9= extremely problematic). The nine-point Likert scale was later grouped into three categories: low (1-3), moderate (4-6), high (7-9).

According to the mean analysis of remedial approach, seven sounds (/æ/, /oʊ/, /ɑ/, /ɔ/, /œ/, /u:/, and /ɔ:/) (M=4.79) were rated moderately and four (/i/, /ɪ/, /e/, /æ/) (M=2.09) were considered slightly problematic in terms of vowels. With regard to consonants, one sound (/θ/) (M=7.55) was highly problematic, three (/ŋ/, /ð/, and /w/) (M=5.27) were moderately problematic, and 20 (/b/, /f/, /d/, /ʃ/, /g/, /h/, /ð/, /k/, /l/, /m/, /n/, /p/, /r/, /s/, /ʃ/, /t/, /v/, /j/, /z/,
and /ʒ/ (M=1.9) were slightly problematic according to the raters. The findings both converge with and diverge from earlier findings. For instance, the vowels considered slightly problematic by the raters were found to be problematic in earlier studies (Hişmanoğlu & Hişmanoğlu, 2011; Ercan, 2018). Likewise, the /ʒ/, /ʤ/, /ʃ/, and /ʧ/ sounds were rated slightly problematic, while they were considered problematic in earlier research (Mahzoun & Han, 2019). On the other hand, the vowels and consonants rated moderately and highly problematic were also reported similarly in previous studies (see the above-cited references).

The raters were also asked if they had any further comments. Rater 3 declared that certain sounds such as /s/ and /z/ could be problematic in terms of perception. Rater 1 and 4 stated that diphthongs were difficult to pronounce and could be problematic. However, they were excluded in this study. Overall, it might be argued that the results concur with the literature since both convergence and divergence exist with regard to the problematicity of the sounds for Turkish teacher trainees, which might result from various contextual and individual factors.

**RQ (2): What English phonemes did the participants fail to perceive with regard to vowels and consonants?**

The participants’ levels of perception of English phonemes were assessed using the POSE test (Shewell, 2004), which addressed the participants 11 vowel and 25 consonant contrasts. Participants were provided with percentages for their correct responses and those phonemic contrasts with a score of at least 75% or greater implies that test-takers generally did not have a problem with those contrasts. It must, however, be noted here that further diagnosis might be required in cases whereby respondents have received 50% in phonemic contrasts with two items. To answer the second research question, the mean percentages for each phonemic contrast were calculated across vowels and consonants, the results of which were provided in Tables 4-5.

**Table 4. Mean percentages for vowel contrasts**

<table>
<thead>
<tr>
<th>Phonemic contrast</th>
<th>Minimal pair example</th>
<th>Functional load (%)</th>
<th>Number of items (f)</th>
<th>Mean percentage (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/ɪ/ - /ɪ/</td>
<td>beet/bit</td>
<td>95</td>
<td>4</td>
<td>94</td>
</tr>
<tr>
<td>/ɪ/ - /ɛ/</td>
<td>bit/bet</td>
<td>54</td>
<td>4</td>
<td>96</td>
</tr>
<tr>
<td>/ɛ/ - /e/</td>
<td>bet/bait</td>
<td>53</td>
<td>4</td>
<td>99</td>
</tr>
<tr>
<td>/ɛ/ - /æ/</td>
<td>bet/bat</td>
<td>51</td>
<td>4</td>
<td>60</td>
</tr>
<tr>
<td>/æ/ - /ɑ/</td>
<td>cat/cot</td>
<td>76</td>
<td>4</td>
<td>97</td>
</tr>
<tr>
<td>/ɑ/ - /ɑr/</td>
<td>cot/cart</td>
<td>31.5</td>
<td>4</td>
<td>90</td>
</tr>
<tr>
<td>/ɑ/ - /ə/</td>
<td>cot/cut</td>
<td>65</td>
<td>4</td>
<td>52</td>
</tr>
<tr>
<td>/ɑ/ - /ow/</td>
<td>cot/coat</td>
<td>N/A*</td>
<td>4</td>
<td>76</td>
</tr>
</tbody>
</table>
A total of 11 phonemic vowel contrasts with a total of 38 items were displayed to the participants via computer and the mean percentages for each contrast were provided in Table 4. The overall mean percentage for vowel discrimination was 84%, which is slightly greater than the 75% threshold as suggested by Shewell (2004). A close look at the results reveals that the participants failed to perceive /ɛ/ - /æ/ and /ɑ/ - /ə/ contrast (M₁=60%, M₂=52%), even /ɑ/ - /ow/ given that the mean percentage (M₃=76%) was slightly higher than the threshold. The /ə/ - /ʊ/ contrast might also be added to the list of phonemes that the participants failed to perceive, considering the number of items (f=2) and mean percentage (M₄=82%) in this phonemic contrast. All in all, it can be argued that four phonemic pairs (eight phonemes) were problematic, to varying degrees, for English teacher candidates. The findings concurred with those of previous studies (Demirezen, 2010, 2017; Hişmanoğlu & Hişmanoğlu, 2011) with regard to such problematic sounds as /ɛ/, /æ/, /ɑ/, /ə/, /ɑ/, and /ʊ/; however, the present study did not match with earlier scholarly work (Hişmanoğlu & Hişmanoğlu, 2011; Ercan, 2018) in that /i/ and /ɪ/ were not found to be problematic in terms of perception.

Table 5. Mean percentages for consonant contrasts

<table>
<thead>
<tr>
<th>Phonemic contrast</th>
<th>Minimal pair example</th>
<th>Functional load (%)</th>
<th>Number of items (f)</th>
<th>Mean percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/p/ - /b/ (initial)</td>
<td>pill/bill</td>
<td>98</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>/p/ - /b/ (final)</td>
<td>cap/cab</td>
<td>14</td>
<td>4</td>
<td>83</td>
</tr>
<tr>
<td>/p/ - /f/ (initial)</td>
<td>pan/fan</td>
<td>77</td>
<td>4</td>
<td>98</td>
</tr>
<tr>
<td>/p/ - /f/ (final)</td>
<td>cup/cuff</td>
<td>17</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>/v/ - /b/ (initial)</td>
<td>vote/boat</td>
<td>29</td>
<td>4</td>
<td>94</td>
</tr>
<tr>
<td>/v/ - /w/ (initial)</td>
<td>vet/wet</td>
<td>22</td>
<td>4</td>
<td>70</td>
</tr>
<tr>
<td>/f/ - /v/ (initial)</td>
<td>fan/van</td>
<td>23</td>
<td>4</td>
<td>99</td>
</tr>
<tr>
<td>/f/ - /θ/ (initial)</td>
<td>free/three</td>
<td>15</td>
<td>4</td>
<td>86</td>
</tr>
<tr>
<td>/θ/ - /t/ (initial)</td>
<td>thin/tin</td>
<td>18</td>
<td>4</td>
<td>76</td>
</tr>
</tbody>
</table>
The participants were provided with 25 different consonant contrasts consisting of 84 items and the mean percentages for each consonant pair were given in Table 5. The overall mean percentage for the consonant section was 94%, which was about 20% greater than the threshold success score. When examined carefully, it might be argued that the respondents failed to perceive /v/ and /w/ (M₁=70%). The number of problematic consonants can be increased if /θ/ - /t/ initial (M₂=76%), /p/ - /b/ final (M₃=83%), /f/ - /θ/ initial (M₄=86%), /ð/ - /d/ initial (M₅=84%), /s/ - /z/ final (M₆=82%), and /k/ - /g/ final (M₇=85%) contrasts are also added to the list due to the little difference between the threshold percentage (75%) and the participants’ actual performances and the insufficiency of number of items (f=2) in certain pairs such as /ð/ - /d/. Overall, it can be argued that seven consonant contrasts (14 consonants) with different phonemic positions were found to be troublesome, to a greater or lesser extent, for prospective English teachers. The findings were in line with those of early studies (Demirezen, 2005b; Hişmanoğlu, 2009) with regard to the problematic consonants; however, other consonants in the test were not found to be tricky in terms of perception, unlike other studies (Ercan, 2018; Mahzoun & Han, 2019) that suggested the problematicity of these sounds.
**RQ (3): How well did the participants perform in terms of accentedness, comprehensibility, and intelligibility?**

The data for the third research questions came from an elicitation paragraph read by the participants. The recordings were evaluated by seven experts using two nine-point Likert-type scales (for accentedness and comprehensibility) and a transcription task (for intelligibility), the results of which were presented in Table 6 and Table 7 consecutively.

**Table 6. Participants’ level of accentedness and comprehensibility**

<table>
<thead>
<tr>
<th>Type of domain</th>
<th>M</th>
<th>N</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accentedness</td>
<td>4.77</td>
<td>49</td>
<td>.941</td>
</tr>
<tr>
<td>Comprehensibility</td>
<td>1.95</td>
<td>49</td>
<td>.351</td>
</tr>
</tbody>
</table>

To interpret the results, the nine-point Likert scale was categorized into three as low (1-3), moderate (4-6), and high (7-9) for accentedness and easy (1-3), medium (4-6), and high (7-9) for comprehensibility. When the mean scores for accentedness (linguistic nativelikeness) and comprehensibility (ease of understanding) are rounded up, it might be stated that the participants were rated moderately accented and their speech was easily understandable by the listeners. This finding diverged from that of Uzun’s (2019) study that found lower comprehensibility in Turkish EFL learners’ speech. The listeners were experienced nonnative teachers of English but inexperienced raters in accentedness and comprehensibility. In this study, the listeners gave higher comprehensibility and moderate accentedness ratings. The listeners can, therefore, be considered lenient in their ratings given that they have language teaching experience (Saito et al., 2017) and familiarity with the speech accent they judge (Winke et al., 2013). It might also be claimed that the raters who were L2 listeners differed in their patterns of ratings from those of L1 listeners, particularly when they rate familiar accents (Foote & Trofimovich, 2018).

With regard to the intelligibility ratings of the listeners, the raters were asked to note down the words they considered unintelligible instead of transcribing the whole text. The productive pronunciation text includes 133 words occurring 306 times. A little below 75% per cent (% = 96) of the words with a frequency of 6577 were marked unintelligible by the listeners. Of the 96 problematic words, one had four-digit frequency (1.04%), 15 had three-digit frequency (15.62%), 58 had two-digit frequency (60.42%), and 22 had one-digit frequency (22.92%) of occurrences as problematic items.
Table 7. High-frequency unintelligible words

<table>
<thead>
<tr>
<th>Words</th>
<th>( f )</th>
<th>Phonemic transcription</th>
<th>Problem sound</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>wolf</td>
<td>1801</td>
<td>/wʊlf/ - /wʊlf/</td>
<td>/ʊ/</td>
<td>word-medial</td>
</tr>
<tr>
<td>shepherd</td>
<td>321</td>
<td>/ˈʃep.ərd/ - /ˈʃep.ərd/</td>
<td>/ə/</td>
<td>word-medial</td>
</tr>
<tr>
<td>hill</td>
<td>228</td>
<td>/hɪl/ - /hɪl/</td>
<td>/ɫ/</td>
<td>word-final</td>
</tr>
<tr>
<td>youth</td>
<td>199</td>
<td>/juːθ/ - /juːθ/</td>
<td>/0/</td>
<td>word-final</td>
</tr>
<tr>
<td>chasing</td>
<td>180</td>
<td>/ˈʃeɪzɪŋ/ - /ˈʃeɪzɪŋ/</td>
<td>/ŋ/</td>
<td>word-final</td>
</tr>
<tr>
<td>comfort</td>
<td>172</td>
<td>/ˈkɑːm.fət/ - /ˈkɑːm.fət/</td>
<td>/ə/</td>
<td>word-medial</td>
</tr>
<tr>
<td>again</td>
<td>152</td>
<td>/ˈɡɛn/ - /ˈɡɛn/</td>
<td>/ə/</td>
<td>word-initial</td>
</tr>
<tr>
<td>grumbling</td>
<td>146</td>
<td>/ˈɡrʌm.blɪŋ/ - /ˈɡrʌm.blɪŋ/</td>
<td>/ŋ/</td>
<td>word-final</td>
</tr>
<tr>
<td>sheep</td>
<td>129</td>
<td>/ʃiːp/ - /ʃiːp/</td>
<td>/iː/</td>
<td>word-medial</td>
</tr>
<tr>
<td>flock</td>
<td>128</td>
<td>/flɑːk/ - /flɒk/</td>
<td>/ɑː/ - /ɒ/</td>
<td>word-medial</td>
</tr>
<tr>
<td>naughty</td>
<td>116</td>
<td>/ˈnɑː.ti/ - /ˈnɔː.ti/</td>
<td>/ə/ - /ɔ/</td>
<td>word-medial</td>
</tr>
<tr>
<td>truth</td>
<td>116</td>
<td>/truːθ/ - /truːθ/</td>
<td>/0/</td>
<td>word-final</td>
</tr>
<tr>
<td>said</td>
<td>113</td>
<td>/sɛd/ - /sɛd/</td>
<td>/ɛ/</td>
<td>word-medial</td>
</tr>
<tr>
<td>alarmed</td>
<td>109</td>
<td>/ˈɑːl.ərm/ - /ˈɑːl.ərm/</td>
<td>/ə/</td>
<td>word-initial</td>
</tr>
<tr>
<td>away</td>
<td>106</td>
<td>/ˈweɪ/ - /ˈweɪ/</td>
<td>/ə/</td>
<td>word-initial</td>
</tr>
<tr>
<td>thought</td>
<td>105</td>
<td>/θɔːt/ - /θɔːt/</td>
<td>/0/</td>
<td>word-initial</td>
</tr>
</tbody>
</table>

A total of 16 words (four in word-initial, seven in word-medial, and five in word-final positions) with a frequency of 4121 (62.65%) was found to be unintelligible by the listeners. Nine sounds (/ʊ/, /ə/, /ɫ/, /θ/, /ŋ/, /iː/, /ɑː/, /ɒ/, /ɔ/, and /ɛ/) were problematic according to the listeners, which concurred with the findings of previous studies (Demirezen, 2010; Hişmanoğlu & Hişmanoğlu, 2011; Bardakçı, 2015; Ercan, 2018; Uzun, 2019, 2021). When analyzed further, 16 words with four- and three-digit frequencies (accounting for 22.60% of text coverage and 62.65% of incorrect word frequency) were determined. Since the focus in this study was only segmentals, suprasegmental features were ignored as regards intelligibility. However, accurate production of sounds in the stressed syllables were deemed significant for intelligibility (Zielinski, 2008). Also, the “international core for phonological intelligibility” (Richards, 2015, p.345) proposed by Jenkins (2000) included certain segmentals that might affect mutual intelligibility in nonnative contexts. Considering the 8417 (56.14%) occurrences...
of correct and 6577 (43.86%) incorrect words in overall 14994 occurrences, it might be argued that accent familiarity did not affect the listeners’ judgments as claimed otherwise (Lane, 2010, p.3).

5. Conclusions and Pedagogical Implications

The present study provided a refreshed perspective on the problematic segmental sounds for Turkish EFL teacher trainees by way of (i) remedial and expert judgment approaches, (ii) diagnostic tests to determine perceptually and productively tricky sounds, and (iii) evaluating their speech in terms of accentedness, comprehensibility, and intelligibility. Expert judgment showed that English segmentals such as /æ/, /oʊ/, /ɑ/, /ə/, /u:, /u/, /e/, /ɛ/, and /ʌ/ tended to be problematic for Turkish EFL teacher trainees, which both accorded and discorded with previous research. This makes sense given that the present research is a case study thus context-dependent. The perception test revealed that vowels such as /ɛ/ - /æ/, /ɑ/ - /ə/ contrast, and /ɑ/ - /ow/ and consonants including /v/ and /w/, /θ/ - /t/ initial, /p/ - /b/ final, /f/ - /θ/ initial, /ð/ - /d/ initial, /s/ - /z/ final, and /k/ - /g/ final were relatively difficult to perceive. The productive test suggested that the participants’ speech was moderately accented and effortlessly comprehensible, and relatively intelligible. It further indicated that /ʊ/, /ə/, /ɫ/, /θ/, /ŋ/, /iː/, /ɑː/, /ɒ/, /ɔ/, and /ɛ/ were the sounds hampering intelligibility at the segmental level. The findings of the perception and productive tests were congruent to remedial approach and expert judgment.

It was clearly demonstrated that certain English sounds persisted to be problematic for Turkish EFL teacher trainees in terms of perception and production that somewhat impedes intelligibility which is key to pronunciation and effective communication (Levis, 2018). This finding has certain pedagogical implications if pronunciation-related problems continue to exist for teacher trainees. Prospective teachers of English are only provided with pronunciation instruction in two undergraduate courses titled “Listening and Pronunciation I/II” in the first year. Given that the second course mainly deals with advanced listening, it is safe to say that there is only one course supplying the trainees with pronunciation instruction. The content of this course does not specifically address to the segmental errors of teacher candidates. It can, therefore, be suggested that this issue should be surmounted, such as by integrating an elective course into the teacher education curriculum that solely deals with remediating problematic sounds for teacher trainees or by reconsidering the course content in relation to these pronunciation errors.

It goes without saying that this study has some limitations. One limitation is the sole focus on vowels and consonants thus the exclusion of diphthongs and suprasegmentals since it aimed to serve a single focus. Another limitation is related to the study’s being a case study notwithstanding the inclusion of both qualitative and quantitative data. However, the findings of this study concurred with those of previous research, contributing to the generalization of the fact that Turkish EFL teacher trainees possessed segmentally challenging English sounds at the perception and production levels. Future studies might consider the inclusion of suprasegmentals and intervention while exploring the problematicity of pronunciation for teacher candidates.
References


1811


