

66. Minimal Pair Perception and the Salience of Phonological Features in Turkish Speakers of English¹

Tarık UZUN²

APA: Uzun, T. (2024). Minimal Pair Perception and the Salience of Phonological Features in Turkish Speakers of English. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (41), 1249-1266. DOI: <https://10.5281/zenodo.13338049>

Abstract

This study explores the sound discrimination responses of Turkish speakers of English through a diagnostic listening test administered in an English as a foreign language (EFL) setting, examining their perceptions of minimal pair (MP) similarities or distinctions. The findings reveal that participants accurately marked a large majority of MPs as the same or different. However, error frequencies varied across items, with some placed in medium to very high error ranges. Errors involving vowels and consonants in these ranges were observed to be evenly distributed, with neither category outweighing the other. It was observed that salient differences among segmentals played a more prominent role in successfully distinguishing MPs in the low error range, while subtle differences, often harder to catch, contributed to higher error rates. Multiple contributing factors including participants' first language (L1) seemed to interact in complex ways and influenced participants' performances of sound discrimination. The study identifies several high functional load (FL) MPs in the medium to very high error ranges, highlighting their crucial role in word differentiation and meaning construction. The results also suggest the need for tailored teaching addressing specific challenges and nuanced distinctions, as well as prioritizing high FL pairs in instruction to enhance learners' communicative competence and intelligibility.

Keywords: Minimal pairs, perception, vowels, consonants, diagnostic test

¹ **Statement (Thesis / Paper):** It is declared that scientific and ethical principles were followed during the preparation process of this study and all the studies utilised are indicated in the bibliography.

Conflict of Interest: No conflict of interest is declared.

Funding: No external funding was used to support this research.

Copyright & Licence: The authors own the copyright of their work published in the journal and their work is published under the CC BY-NC 4.0 licence.

Ethics Statement: Ethical permission was granted by the Ethics Commission of Ankara Yıldırım Beyazıt University with the decision dated 01.03.2023 and numbered 37.

Source: It is declared that scientific and ethical principles were followed during the preparation of this study and all the studies used are stated in the bibliography.

Similarity Report: Received - Turnitin, Rate: %6

Ethics Complaint: editor@rumelide.com

Article Type: Research article, **Article Registration Date:** 06.02.2024-**Acceptance Date:** 20.08.2024-

Publication Date: 21.08.2024; DOI: <https://10.5281/zenodo.13338049>

Peer Review: Two External Referees / Double Blind

² Öğr. Gör. Dr., Ankara Yıldırım Beyazıt Üniversitesi, Yabancı Diller Yüksekokulu / Lect. Dr., Ankara Yıldırım Beyazıt University, School of Foreign Languages (Ankara, Türkiye), uzuntarik@aybu.edu.tr, **ORCID ID:** <https://orcid.org/0000-0002-9243-963X> **ROR ID:** <https://ror.org/05ryem72>, **ISNI:** 0000 0004 0454 9762 **Crossref Funder ID** 501100014801

Türk İngilizce Konuşucularında En Küçük Çiftlerin Algılanması ve Sesbilimsel Özelliklerin Belirginliği³

Öz

Bu çalışma, Türk İngilizce konuşurlarının ses ayırt etme yanıtlarını araştırmak amacıyla, yabancı dil olarak İngilizce bağlamında uygulanan bir tanılayıcı dinleme testi yardımıyla en küçük çift benzerlikleri ve ayrımlarına yönelik algıları incelemektedir. Bulgular, testte sunulan en küçük çiftlerin büyük bir bölümünün doğru bir şekilde aynı veya farklı olarak işaretlenebildiğini göstermektedir. Bununla birlikte, hata frekansları öğeler arasında değişiklik göstermiş, hatalı yanıt verilen maddelerin bir bölümü orta ila çok yüksek hata aralıklarına yerleşmiştir. Bu aralıklardaki ünlü ve ünsüz hatalarının, eşit bir dağılım gösterdikleri gözlenmiştir. Parçalı sesbirimler arasında gözlenen belirgin farkların, özellikle düşük hata aralığındaki en küçük çiftleri başarıyla ayırt etmede daha etkili olduğu, genellikle algılanması güç ve daha zor yakalanabilen farkların daha yüksek hata oranlarına neden olduğu anlaşılmıştır. Ses ayırt etme performanslarında, katılımcıların birinci dili de dahil olmak üzere çeşitli etmenlerin karmaşık etkileşimlerinin yansımaları gözlenmiştir. Orta ila çok yüksek hata aralıklarında yüksek işlevsel yüke sahip çok sayıda en küçük çiftin varlığı tespit edilmiş, bu çiftlerin sözcük ayrımı ve anlam oluşturmadaki kritik rolleri üzerinde durulmuştur. Sonuçlar, yaşanan zorlukları ve sesler arasındaki ince ayrımlara odaklı bir öğretime duyulan gereksinimi ve öğrenenlerin iletişimsel yeterlikleri ile anlaşılabilirliklerini geliştirmek üzere yüksek işlevsel yüklü en küçük çiftlere öğretimde öncelik verilmesi gerekliliğini ortaya koymuştur.

Anahtar Kelimeler: En küçük çiftler, algı, ünlüler, ünsüzler, tanısal test

³ **Beyan (Tez/ Bildiri):** Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.

Çıkar Çatışması: Çıkar çatışması beyan edilmemiştir.

Finansman: Bu araştırmayı desteklemek için dış fon kullanılmamıştır.

Telif Hakkı & Lisans: Yazarlar dergide yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmaları CC BY-NC 4.0 lisansı altında yayımlanmaktadır.

Etik İzni: Ankara Yıldırım Beyazıt Üniversitesi Etik Komisyonu tarafından 01.03.2023 tarihli, 37 sayılı kararla etik izni verilmiştir.

Kaynak: Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.

Benzerlik Raporu: Alındı – Turnitin, Oran: %6

Etik Şikayeti: editor@rumelide.com

Makale Türü: Araştırma makalesi, **Makale Kayıt Tarihi:** 06.02.2024-**Kabul Tarihi:** 20.08.2024-**Yayın Tarihi:** 21.08.2024; **DOI:** <https://10.5281/zenodo.13338049>

Hakem Değerlendirmesi: İki Dış Hakem / Çift Taraflı Körleme

1. Introduction

Pronunciation is an integral aspect of spoken language. The success of comprehension and intelligible pronunciation in a second language (L2) depends on the correct perception and differentiation of its sounds, marking the transition from recognizing unfamiliar sounds to meaningful interpretation (Przedlacka, 2018). In this context, MPs hold considerable importance for instruction and learning. Diagnostic assessment based on MPs is commonly utilized in determining learners' skills to differentiate consonant and vowel sounds (Celce-Murcia et al., 2017) and it helps shape further instructional decisions and actions to take. Isbell (2020) emphasizes the need for a systematic diagnostic evaluation to identify the weaknesses of students, enabling teachers and learners to give priority to specific areas in learning and teaching activities.

Research on minimal pair perception in English among Turkish learners remains notably scarce in the related literature. To address this research gap, this study examined the performance of Turkish speakers of English on a diagnostic sound discrimination test with a particular focus on contributing factors, including the study participants' L1, that potentially impact MP perceptions. With these considerations, the following research questions were answered in the study:

1. How do listeners perform overall on a diagnostic sound discrimination test across MPs?
2. Which phonological features do listeners rely on in sound discrimination across errors with different frequencies?
3. What is the distribution of FL values among the MPs that cause perception difficulties?
4. How do Turkish listeners' L1 affect their perceptions of MPs?

2. Literature review

Understanding and producing speech sounds, involving both perception and production, are fundamental aspects of acquiring spoken proficiency in an L2. These concepts are intricately linked to listening and speaking skills, particularly in the context of second language acquisition (SLA). One effective way to enhance these skills is through the use of MPs. In this regard, MPs defined as "*words or utterances which differ only by one phoneme*" (Kelly, 2012, p. 18) play a critical role in the perception of L2 speech accurately. For instance, the words *hear* and *fear* are two words distinguished by a consonant (/h/ - /f/) changing the meanings in the given examples. Barlow and Gierut (2002) emphasize that the distinctions in MP contrasts can vary from minimal to maximal differences. They further explain that a minimal contrast entails only a few feature distinctions among phonemes, whereas a maximal contrast signifies a phonemic difference that extends across multiple dimensions, including place, manner, and voice.

In addition to their theoretical usefulness, MPs have also played an important role in pronunciation instruction, being utilized for diagnostic evaluations, oral practice, and listening comprehension materials (Levis & Cortes, 2008). More specifically, they are applied in various pronunciation tasks, including identification, discrimination, isolation, sorting, sentence discrimination, and picture tasks (Avery & Erlich, 1992). Several studies in different contexts investigated the effectiveness of using MPs in instruction and reported positive results (e.g., Bradlow et al., 1999; Haghighi & Rahimy, 2017; Putra & Rochsantiningsih, 2018; Valenzuela & French, 2023).

Linguistic inquiries and pedagogical interventions into MPs and L2 perception need to take learners' L1 into consideration as a variable. Flege (1988) suggests that due to cross-linguistic differences in sound realization, only a limited number of sounds in an L2 are produced identically to those in L1 and it is probable that learners will perceive certain sounds in L2 differently from native speakers. Building upon his earlier work, Flege (2003) provides a more nuanced perspective on the inherent challenges of L2 vowel perception acknowledging the inherent mismatch between L1 and L2 vowels even if they have similar values. Highlighting the problematic nature of vowels in English for learners in producing speech, Levis & Barriuso (2012) remind that there are 5 vowel letters, and 15 vowel sounds in English resulting in a very indirect sound spelling correspondence for learners, which might lead to difficulties for L2 learners.

2.1. Segmental challenges for Turkish learners of English

Several studies have discussed segmental difficulties or errors of learners with a Turkish L1 background in English. Research focusing on preservice English teachers (PrETs) is particularly common in this area. For example, Arıkan and Yılmaz (2020) examined the pronunciation errors among PrETs, revealing issues such as vowel shortening, insertion, devoicing of final consonants, and gemination. They also identified specific segmental errors, including the mispronunciations of /ŋ/, /ð/, /θ/, /w/, /ɾ/, /oo/, /æ/, and /□/. Bardakçı (2015) also explored PrETs' common pronunciation errors and identified /□/, /θ/, /ŋ/ and /æ/. Besides these, /æ/ and /ʌ/ (Demirezen, 2008), /□/ (Demirezen, 2010), and /v/ - /w/ (Demirezen, 2005) were named as fossilized pronunciation errors for Turkish speakers of English.

Uzun (2022) explored preservice English teachers' salient pronunciation errors based on expert listener judgments and their relative importance on their intelligibility in English via naive listener transcriptions of mispronounced target words. The mispronunciations of /□/ and /θ/ emerged as the most prevalent segmental errors, yet they did not noticeably compromise intelligibility. In a more recent study, Gökgöz-Kurt (2023) examined how native Turkish speakers perceive the bilabial glide /w/ in L2 English and found accurate perception in discrimination and identification tasks, with higher success in identifying phonemes that are not present as allophones in L1 Turkish during the identification task.

2.2. Functional load hypothesis and varying strengths of minimal pairs

King (1967) defines FL as a measure of the number of MPs identifiable for a particular phonemic opposition serving as an indication of the role played by two phonemes (or a distinctive feature) in differentiating utterances. As per Catford (1987), the FL of a phoneme or phonemic contrast is expressed through the count of words featuring it in the lexicon, or by the number of words in the lexicon where it contributes to maintaining distinctions. Derwing and Munro (2015) demonstrated, for instance, that the FL of the /p/-/b/ contrast in English has a high FL, given the abundance of MPs separated by these sounds (e.g., *pea* vs. *bee*, *pat* vs. *bad*), in contrast to the /θ/-/ð/ pair, which distinguishes a comparatively more limited set of words in the language (e.g., *thigh* vs. *thy*).

Catford (1987) and Brown (1988) presented FL values of vowel and consonant MPs within the form of relative rank orderings in their works. In Catford's (1987) list of *Relative Functional Load*, MPs were listed as initial consonants, final consonants, and vowels, and ranked as percentages (higher percentages representing higher FLs). In Brown's (1988) *Rank Ordering of RP Phoneme Pairs Commonly Conflated by Learners*, where vowel and consonant pairs were ranked on a scale from 1 to 10, a ranking of 1 indicated *very low FL*, while a ranking of 10 signified *very high FL*. Levis (2018) and Munro and

Derwing (2006) emphasize the importance of high FL in achieving intelligible and comprehensible speech, reinforcing the need to prioritize these aspects in instruction.

2.3. Diagnostic assessment of pronunciation

An essential aspect of instructional planning involves conducting needs assessment to guide the instructional direction. Derwing and Munro (2015) emphasize that despite the existence of curricula in many language programs, a needs assessment is crucial for tailoring instruction to the unique requirements of individual learners. In this regard, diagnostic tests serve as valuable tools for educators as they are specifically designed to pinpoint learners' specific strengths and weaknesses in various language aspects (Knoch, 2017).

Regarding pronunciation instruction, a crucial initial step for teachers is to assess each individual's pronunciation, a practice recommended not only for this specific area, but also in other educational domains (Derwing, 2023). A diagnosis of this sort can be done in several ways, such as conducting listening comprehension tests, evaluating free speech, or assessing reading aloud abilities at different linguistic levels, including word, sentence, or discourse levels (Levis & Barriuso, 2012). Diagnostic assessment, in this context, can focus on the perception or production aspects of learners' speech through varied tasks and designs. In terms of perception, instructional settings often employ diagnostic listening tests that involve specific sound discrimination tasks. These tasks are commonly featured in educational resources and can assess the learner's capacity to differentiate between segmental and suprasegmental speech features, including vowel and consonant sounds, word stress, and intonation (Celce-Murcia et al., 2017; Derwing & Munro, 2015; Kang & Kermad, 2018).

Teachers are the first-hand users of the data accessed via diagnostic perception tests as they are to utilize them for their own instructional decisions and implementations. According to Kang and Kermad (2018), teachers' roles in the process of ongoing assessment based on the diagnostic measures and needs assessment analysis are a) to evaluate students' progress so as to individualize the curriculum and b) to provide individual and continuous feedback on their progress.

As Derwing (2023) highlights, accurate perception of both segmental and prosodic features in the L2 is closely linked to learners' production abilities, and if learners struggle to differentiate phonologically important features, it means a focused attention on perception skills is needed. In a study conducted by Kissling (2014), learners received explicit pronunciation instruction through modules integrated into their curriculum of Spanish as a foreign language, followed by assessments of their pronunciation accuracy. According to the results of the study, instructors are advised to allocate sufficient time for learners to develop their perceptions of target speech sounds at the beginning of pronunciation teaching, as the initial proficiency in perceiving these sounds partly determines the effectiveness of the instruction and subsequent learning outcomes.

3. Method

3.1. Data collection

3.1.1. Participants

This quantitative research took place at a state university's school of foreign languages in Türkiye. Participants (N=100) were students enrolled in B1 level classes of the pre-faculty mandatory English

preparatory program in 2023-2024 academic year over the course of four months in different groups. In the preparatory program, students are placed in B1 level classes and higher levels based on their performance in an in-house English proficiency exam that assesses all four language skills, namely reading, writing, listening, and speaking. All students enrolled in the preparatory school attend skill-based classes and must pass the same proficiency exam administered by the school to complete the program successfully. Table 1 summarizes the participant demographics:

Gender	N	Age Group	N
Male	39	17-25	94
Female	61	26-35	6
Total: 100			

Table 1. Participant Demographics

As the demographic data suggest, the majority of the participants were female (%61) and between the ages 17-25 (%94) at the time of data collection. All the participants in the study were from Türkiye and were the native speakers of Turkish. The majors of the participants included diverse disciplines such as Law, Medicine, Engineering, and social sciences. Students majoring in languages (i.e., English Translation and Interpretation program) were excluded from the study. Participation was voluntary, and all participants provided informed consent before completing the demographics questionnaire. Following these procedures, a diagnostic perception test was administered to participants in their respective classes under the researcher's supervision.

The diagnostic test was administered in classrooms equipped with high-quality speakers. The researcher provided the necessary instructions and guided participants through two training items in the same format as the actual test. After ensuring that the participants understood the task, the audio input was initiated. Participants were given the opportunity to listen to each pair once, as presented in the audio recording. Under the guidance of the researcher, participants were instructed to respond to each item by marking their answers on the provided answer sheet using pen and paper as they listened. The length of the recording was 5.08 minutes.

This study was granted ethical approval by the Ankara Yıldırım Beyazıt University Ethics Committee on March 15, 2023, under approval number 03-37.

3.1.2. Instrument

This study utilized Baker's (2006) diagnostic perception test to assess participants' proficiency in segmental and suprasegmental speech features at the beginning of dedicated in-class pronunciation interventions across four classes during a semester. The test comprised three sections focusing on distinct aspects of spoken language:

- Section I - Sound Discrimination (51 items) assessed participants' ability to differentiate between pairs of words through auditory perception. For each pair, they had to mark whether the words were the *same* or *different* on their answer sheet.
- Section II - Intonation (10 items) evaluated participants' ability to perceive final intonation patterns in the utterances they heard. Upon hearing each utterance, they marked whether the final intonation *rose* or *fell* at the end of each utterance.

- Section III - Word Stress (5 items) tested participants' perception of word stress placement. Each set contained five words, one differing in stress pattern. Their task was to identify the word with the different stress placement.

While the entire test was administered, only the results from *Section I - Sound Discrimination* were used for subsequent analysis within the scope of this study. This section contained 51 items, comprising 10 pairs presented with the same pair of words (e.g., *ill / ill*), and 41 pairs featuring distinct variations in terms the MPs featured (e.g., *tin / ten*). Among these items with distinct MPs, 20 involved consonant differences (e.g., *ship / chip*) and 20 pairs contrasted vowel sounds (e.g., *bed / bird*). Besides these, one pair (*big / bigger*) presented a unique case, featuring an extra syllable in the second word. All items were pronounced in *British English* (BE) with its speech characteristics including non-rhoticity, in other words, the lack of /r/ sound in coda positions in Received Pronunciation (RP, or BE) and several other varieties of the language (Carr, 2008). In BE, the phoneme /r/ is typically pronounced only before a vowel (pre-vocalic position), as in words like *red* and *arrive*, but not after a vowel (post-vocalic position), as in words like *hear* and *bird* (Rogerson-Revell, 2018, p.107).

3.2. Data analysis

The participants' performance on the diagnostic test was analyzed by the researcher using descriptive statistics, including frequencies and percentages of correct and incorrect responses based on the answer key. Analyses were performed using Jamovi (2022), a software for a statistical analysis. For each item, the number of incorrect answers was calculated and used to categorize the items into four error ranges:

- **Low error range:** Less than 10% of the participants answered incorrectly. These items likely represent skills most participants demonstrated strong proficiency.
- **Medium error range:** 10% to 29% of the participants answered incorrectly. This range may indicate items where a sizable portion of participants encountered some difficulty, though most still answered correctly.
- **High error range:** 30% to 49% of the participants answered incorrectly. This range suggests items where a considerable number of participants struggled.
- **Very high error range:** 50% or more of the participants answered incorrectly. This range identifies items where the majority of participants had difficulty.

The categorization of error ranges was designed by the researcher based on the observed distribution of errors. Items with less than 10% incorrect answers were classified as *low error range*, reflecting their prevalence. The *medium* and *high error ranges* were established according to observed distributions and significant gaps. The *very high error range* was assigned to items with errors exceeding 50%, though these instances were limited in number. This approach ensured clear differentiation in difficulty levels and error patterns within the specific dataset of this study. Based on participants' performances, items that fell into *medium* to *very high* error ranges were further analyzed and classified for their phonological properties. Items in low error range were also examined as a complementary source of data.

4. Results

4.1. Distribution of items into error ranges

To analyze participant performance on the test, items were first categorized based on the frequency of errors made in each, resulting in four distinct error ranges as explained in the method: *low*, *medium*, *high*, and *very high*. In response to the first research question, Table 2 presents participants' overall performance in the diagnostic perception test administered, detailing the numbers and percentages of items within each error range:

Ranges	N	%
Low Error Range	32	62,74
Medium Error Range	12	23,52
High Error Range	4	7,85
Very High Error Range	3	5,89
Total	51	100

Table 2. Distribution of items into error ranges

The distribution of error percentages revealed a clear majority of items (%62,74) falling into the *low error range*, suggesting that fewer than 10 participants made mistakes in these items. The remaining 19 items (%37,26) spread across the other three ranges, with *medium* being the densest (%23,52). A combined %13,74 of items fell within the *high* and *very high* error ranges, indicating considerable challenges for most participants.

4.2. Error types in medium to very high error ranges

In response to the second research question, a detailed examination was conducted to determine which phonological features listeners relied on for sound discrimination across different error ranges. Overall, there is a balanced distribution of challenges across 10 vowel and 9 consonant MPs in the *medium-to-high* error ranges. Neither category seemed to pose notably more trouble than the other. Vowel and consonant errors are evenly distributed with 10 vowel and 9 consonant errors. As also shown previously in Table 2, the medium error range remained the most populous. Table 3 presents the comprehensive results of the examination of the specific MPs that fell within *medium* to *very high* error ranges:

Item No.	Item	Error Range	Type	Number of Errors	MP
1b	ship - chip	High	Consonant	32	/ʃ/-/tʃ/
2a	six - seeks	Medium	Vowel	14	/ɪ/-/i:/
3a	tin - ten	Medium	Vowel	13	/ɪ/-/ε/
4a	said - sad	High	Vowel	44	/ε/-/æ/
4b	choke - joke	High	Consonant	45	/tʃ/-/dʒ/
5b	much - march	Very high	Vowel	55	/ʌ/-/ɑ:/
6b	few - view	Medium	Consonant	13	/f/-/v/
7c	thing - thin	Very high	Consonant	71	/ŋ/-/n/
8a	sport - spot	Medium	Vowel	16	/ɔ:-/ɒ/

9a	put - pot	Medium	Vowel	12	/ʊ/-/ʊ/
10b	grass - glass	Medium	Consonant	11	/r/-/l/
11a	cup - cap	Medium	Vowel	23	/ʌ/-/æ/
12a	comb - cone	Very high	Consonant	84	/m/-/n/
13b	closing - clothing	Medium	Consonant	10	/z/-/ð/
14a	girl - gull	Medium	Vowel	15	/ɜ:/-/ʌ/
14d	drain - train	Medium	Consonant	13	/d/-/t/
16b	fern - phone	High	Vowel	30	/ɜ:/-/əʊ/
17a	laugh - life	Medium	Vowel	12	/ɑ:/-/aɪ/
23b	free - three	Medium	Consonant	22	/f/-/θ/

Table 3. Breakdown of target MP characteristics and error frequencies

Even though the instances are not many, Table 3 also presents three items within the very high error each with over 50 occurrences. Two pairs within this range exceeded 70 occurrences, indicating the highest level of misidentification among the participants.

4.3. Medium error range MPs

A closer examination in medium error range MPs (see Table 4) indicate that vowel and consonant pairs present certain characteristics. Most vowel pairs involve differences in tongue position in vowels (closeness – openness), frontness/backness (front – central – back), or diphthong presence (monophthong vs. diphthong). The consonant pairs include voicing distinctions (voiced vs. unvoiced), and place of articulation differences (e.g., alveolar vs. dental). An important finding was that these differences were not salient all the time, yet subtle differences were more common.

Item No.	Item	Number of Errors	MP
2a	six - seeks	14	/ɪ/-/i:/
3a	tin - ten	13	/ɪ/-/ɛ/
8a	sport - spot	16	/ɔ:/-/ʊ/
9a	put - pot	12	/ʊ/-/ʊ/
11a	cup - cap	23	/ʌ/-/æ/
14a	girl - gull	15	/ɜ:/-/ʌ/
17a	laugh - life	12	/ɑ:/-/aɪ/
6b	few - view	13	/f/-/v/
10b	grass - glass	11	/r/-/l/
13b	closing - clothing	10	/z/-/ð/
14d	drain - train	13	/d/-/t/
23b	free - three	22	/f/-/θ/

Table 4. Items in medium error range

More specifically, analyses into vowel pairs within medium error range reveals several factors contributing to difficulties:

- **Tongue position and frontness / backness contrasts:** The vowel pairs in this range show both salient and subtle differences in tongue height, tongue advancement, and rounding, contributing to their distinct phonetic qualities. For instance, in the pair /ɪ/-/i:/, both vowels share a close tongue height. Pairs like /ʌ/-/æ/ and /ɜ:/-/ʌ/ involve subtle distinctions in both tongue height and advancement.

Most vowel pairs belong to the same frontness/backness category (i.e., front-front, central-central, back-back). For instance, /ɪ/-/i:/ and /ʌ/-/ε/ are front vowels and /ɔ:/-/ɒ/ are back vowels. In such cases, the lack of frontness/backness contrasts might be expected to contribute to listener difficulty since these vowel contrasts have less distinct acoustic cues compared to more extreme distinctions such as front vs. back vowels.

- **Tenseness / laxness contrasts:** Tense vowels overall exhibit increased muscular tension, extended duration, and greater subglottal air pressure, while lax vowels are marked by an overall muscular relaxation, minimal articulatory movements, low-pressure air flow, and vowel centralization (Trask, 1996). Three pairs (i.e., /ɪ/-/i:/, /ɔ:/-/ɒ/, /ɜ:/-/ʌ/) present a tense vowel vs. a lax vowel, and in one (item 17a; /ɑ:/-/aɪ/), a tense vowel is contrasted with a diphthong. It is important to note that learners might have encountered challenges in perceiving tense vs. lax distinctions in English, as the prevalence of such distinctions in their L1 phonological systems was different. Additionally, the presence of a diphthong in one pair adds an additional feature of complexity, potentially causing learners to struggle with the unfamiliarity of a comparison between a tense vowel and a diphthong.

In terms of consonants, place of articulation, manner of articulation, and voicing stand out as important features that define the errors in medium error range.

- **Place of articulation and the role of consonant clusters:** The /r/-/l/ pair (as in *grass* vs. *glass*) exhibits proximity in place of articulation, where /r/ is a postalveolar approximant and /l/ is an alveolar lateral approximant. Notably, 11 participants failed to distinguish words in this particular item. In another item (*road* vs. *load*), contrasting /r/-/l/ in the onset position, all participants successfully differentiated the words. This adds to the possibility that comparing individual sounds in consonant clusters might present challenges for listeners in distinguishing consonant pairs. Similar examples support this proposal. For instance, the /f/-/θ/ pair (as in *free* vs. *three*) and /d/-/t/ (as in *drain* vs. *train*) found in this error range are both observed within consonant clusters that exhibit a similar structure, each followed by /r/. This shared coarticulation and phonemic similarity may have contributed to difficulties in differentiation.
- **Manner of articulation:** All consonant pairs share the same manner of articulation with three pairs being fricative (/f/-/v/, /z/-/ð/, /f/-/θ/), one pair plosive (/d/-/t/), and one liquid (/r/-/l/). Similar to place of articulation, subtleties in manner of articulation appear to have made it difficult for listeners to distinguish the MPs in the consonants in question.
- **Voicing distinctions:** The role of voicing, in addition to manner and place of articulation, emerged as another factor influencing participants' perceptions. Among the five consonant pairs, three exhibit a voicing distinction (/f/-/v/, /d/-/t/, /r/-/l/), leading participants to encounter a moderate level of difficulty in distinguishing the voiced and unvoiced sounds within each pair. However, it's important to approach the /r/-/l/ distinction with caution. In BE, the non-rhotic /r/ is not voiced, often realized as silent. Since the audio input was in BE, and the occurrences of /r/ were non-rhotic, they were considered unvoiced in this study. This handling

would have differed in General American English.

4.4. High and very high error ranges

Given the limited number of items in the *high* and *very high* error ranges (error frequency ≥ 30), these ranges were combined for further analysis, resulting in a total of seven items (three vowel and four consonant pairs) falling into this category. A substantial proportion of the participants struggled to differentiate the word pairs, marking them as the *same*, which were actually different. In some cases, this failure extended to over 50% of the total number of participants, indicating a majority of learners encountering difficulties. The specific word pairs and the associated MPs are demonstrated with their frequencies in Table 5:

Item No.	Item	Number of Errors	MP
4a	said - sad	44	/ɛ/-/æ/
16b	fern - phone	30	/ɜ:/-/əʊ/
5b	much - march	55	/ʌ/-/ɑ:/
1b	ship - chip	32	/ʃ/-/tʃ/
4b	choke - joke	45	/tʃ/-/dʒ/
7c	thing - thin	71	/ŋ/-/n/
12a	comb - cone	84	/m/-/n/

Table 5. Items in high and very high error ranges

Close examinations into these errors reveal more similarities and nuanced distinctions rather than clear contrasts. In this context, the resemblances in height and frontness/backness observed in the MPs seem to contribute to perception challenges.

- **Height and frontness/backness:** The vowel pairs have similar tongue positions and frontness/backness. In the case of /ɛ/-/æ/ (as in **said vs. sad**), both vowels are front, however, /ɛ/ has a slightly higher tongue position than /æ/. In the /ʌ/-/ɑ:/ pair, /ʌ/ is a mid-central or mid-back vowel while /ɑ:/ is an open back vowel emphasizing slight differences again in tongue position and frontness/backness.

It is important to handle vowel qualities in relation to consonantal properties at this point. Considering the /ʌ/-/ɑ:/ pair, for instance, the non-rhoticity observed in the word *march*, where the /r/ is not pronounced, likely added a layer of difficulty for listeners. Similarly, in item 16b, where *fern* and *phone* are compared, a similar difficulty appears to have emerged. The initial word *fern* contains an unpronounced /r/ sound, and that the absence of this sound may have led participants to perceive the words in the pair as not distinct.

Regarding the consonants within this range of errors, the overall situation becomes more intricate. The complexity is further compounded by the restricted number of relevant MPs.

- **Place and manner of articulation and voicing:** In the given consonant pairs (/ʃ/-/tʃ/, /tʃ/-/dʒ/, /ŋ/-/n/, and /m/-/n/), what stands out is the diverse set of phonological patterns and distinctions. The pairs showcase variations in voicing, manner of articulation, and specific

places of articulation. For example, the contrast between /ʃ/ and /tʃ/ (as in **ship vs. chip**) highlights a shift from a voiceless postalveolar fricative to a voiceless postalveolar affricate. Similarly, the /tʃ/-/dʒ/ (as in **choke vs. joke**) pair demonstrates a change in voicing within the postalveolar affricate category. Moving to /ŋ/-/n/ (as in **thing vs. thin**), the contrast involves the shift from a velar nasal to an alveolar nasal, emphasizing differences in the place of articulation. Lastly, the /m/-/n/ (as in **comb vs. cone**) pair demonstrates distinctions in place of articulation, with /m/ being a bilabial nasal and /n/ an alveolar nasal.

4.5. MP perception errors and functional load hypothesis

In response to the third research question, the MPs identified within medium to very high error ranges were further analyzed in relation to their FLs based on Brown (1998), aiming to assess their relative importance within instances of conflation. The selection of this list was motivated by its alignment with BE utilized in the audio input for this study. The detailed outcomes of this assessment are presented in Table 6:

Item No.	Item	MP	FL Value
1b	ship - chip	/ʃ/-/tʃ/	2
2a	six - seeks	/ɪ/-/i:/	8
3a	tin - ten	/ɪ/-/ε/	9
4a	said - sad	/ε/-/æ/	10
4b	choke - joke	/tʃ/-/dʒ/	3
5b	much - march	/ʌ/-/ɑ:/	5
6b	few - view	/f/-/v/	7
7c	thing - thin	/ŋ/-/n/	5
8a	sport - spot	/ɔ:/-/v/	5
9a	put - pot	/ʊ/-/v/	NA*
10b	grass - glass	/r/-/l/	10
11a	cup - cap	/ʌ/-/æ/	10
12a	comb - cone	/m/-/n/	10
13b	closing - clothing	/z/-/ð/	7
14a	girl - gull	/ɜ:/-/ʌ/	5
14d	drain - train	/d/-/t/	9
16b	fern - phone	/ɜ:/-/əʊ/	9
17a	laugh - life	/ɑ:/-/aɪ/	9
23b	free - three	/f/-/θ/	1

*The MP is not included in Brown's (1988) list.

Table 6. FL Values of Medium to Very High Range Errors

The numerical values reveal that the majority of the examined MPs exhibit high FLs, 11 pairs falling within the range of FL ranks 7 to 10, and an additional 4 pairs carry a FL of 5 to 7. In contrast, only 3 pairs were identified with low FLs. To put it another way, most of the MPs that participants failed to distinguish carried high importance in meaning differentiation, which could potentially confuse them

and easily lead to misunderstandings.

4.6. MP perception errors and participants' L1

Regarding the fourth research question, the difficulties faced by Turkish listeners in discerning specific English sounds arise from a combination of factors, with potential L1 influence being a contributor. This is primarily due to the absence of certain phonological distinctions in Turkish. One noteworthy phonological feature that emerged as a distinctive element in participants' errors within the medium to very high range was the distinction between tense and lax vowels, also defined in terms of vowel length. All vowels in the Turkish sound system are generally lax or short, with some exceptions that can be lengthened in specific words (Ergenç, 2002). Among the MPs examined, four involved a contrast between a tense vowel and a lax vowel, while two others featured a tense vowel versus a diphthong.

The absence of rhoticity in BE may have contributed to participants' difficulties in sound discrimination, as English tends to drop the /r/ sound in non-rhotic varieties, presenting a contrast with Turkish, which employs a distinct articulation for the /r/ sound. In Turkish, /r/ is a flapped, voiced consonant characterized by taps onto the alveolar ridge using the tip of the tongue; however, when occurring as the final sound in a word, it transforms into a fricative, losing its tapping and voicing features (Ergenç & Bekar Uzun, 2017). Notably, in four MPs within the medium to very high error ranges, the /r/ sound was not pronounced due to the non-rhoticity of English (e.g., *much* vs. *march*, *fern* vs. *phone*), and this feature of BE may have adversely impacted the precise discrimination.

These features align with Flege's (1995) Speech Learning Model (SLM) to a certain extent. An argument developed in SLM is that increased perceived phonetic dissimilarity between an L2 sound and its closest L1 sound enhances the possibility of discerning phonetic differences between them. Turkish lacks a clear tense-lax distinction, meaning learners lack native categories for these subtle length differences in English vowels. This considerably reduces the perceived phonetic dissimilarity between the pairs, possibly making them harder to distinguish. Besides this, the non-existence of rhoticity in BE can lead to confusion, especially when other phonetic differences are subtle. Additionally, the non-existence of certain English sounds in standard Turkish (e.g., /ð/, /ŋ/, /æ/) or specific phonological categories like tense vowels appear to pose a challenge for some listeners. When confronted with unfamiliar sounds, learners tend to assimilate them into their existing L1 categories, which is also in line with the Perceptual Assimilation Model (PAM) (Best, et al., 1988).

4.7. A brief look at low range MP errors

Although not central to the study, a brief examination of MPs that participants distinguished successfully revealed additional insights which also support the previous findings more indirectly. Participants consistently identified 10 specific pairs without difficulty, often because these MPs had more distinct characteristics. For example, in item 13a (*big* vs. *bigger*), the additional syllable in the second word and in 11b (*heel* vs. *ill*), the extra consonant in the initial position made the distinction clearer. Furthermore, among the 10 vowel pairs contrasted, five involved a monophthong and a diphthong, which provided distinct acoustic cues for differentiation. Pairs like /eə/ vs. /ɪə/ (in *air* vs. *ear*) and others with distinct places of articulation (e.g., /ɔ:/ vs. /æ/ in *caught* vs. *cat*) were easier for participants to distinguish. Similar patterns were observed with consonants, where clear differences in articulation, such as /f/ (labiodental) vs. /h/ (glottal) or /b/ (plosive) vs. /v/ (fricative), facilitated accurate identification.

5. Conclusion

This study examined the sound discrimination responses of listeners in a diagnostic test, aiming to investigate their perceptions of MP similarities or distinctions. An important finding was that participants responded to a large majority of the items accurately marking MPs same or different successfully. All the items containing same word pairs were answered correctly by large majorities of participants, yet frequencies of errors differed across items, and they were placed in medium to very high error density ranges. Overall, fewer items were placed in medium to very high error ranges than the items in the low error range maintaining a balance between vowel and consonant errors.

A closer look at the items in low error range items and higher range errors revealed a main difference. In general, salient differences played a more prominent role in successfully distinguished MPs. On the other hand, in most higher error range items, the differences were subtle in terms of phonological features. In other words, similarities rather than contrasts made it harder for listeners to successfully differentiate words with MPs. In addition, in several cases, not a single speech feature but a combination of multiple contributing factors interacting in complicated ways seemed to influence the accuracy of participant responses. L1 influence seems to motivate perception errors, and some of these errors can be interpreted in terms of perceived phonetic dissimilarity, as posited by Flege's (1995) SLM, and assimilation into existing L1 categories, as underlined in Best et al.'s (1988) PAM.

A majority of the MPs identified in the medium to very high error ranges exhibited high FLs, indicating their substantial role in differentiating words and constructing meaning in the language. This result underlines the importance of accurately perceiving these sound pairs, as misinterpretations could potentially result in misunderstandings and contribute to breakdowns in communication. Several studies reach similar results that incorrect high FL substitutions negatively affect comprehensibility (Alnafisah, et al., 2022; Munro & Derwing, 2006; Suzukida & Saito, 2021).

The findings of this study have certain pedagogical implications for EFL instruction. The varying error frequencies across items, ranging from medium to very high, suggest that certain MPs pose greater challenges for learners. The identified subtle differences in higher error range items highlight the need for targeted training to enhance learners' sensitivity to nuanced phonological distinctions. Educators might consider prioritizing the misperceived pairs falling into the medium to very high error ranges in instruction to ensure learners develop proficiency in discriminating them. High FL MPs should also be targeted in perception and production to improve learners' communication skills and clarity.

Despite the insights it has provided, this study also has some limitations. Primarily, the focus was confined to segmentals, and suprasegmental speech features such as word stress, intonation, and sentence stress were not included in the investigation. To address this need, Uzun (2024) investigated EFL learners' final intonation (rise/fall) and word stress placement in a follow-up study. Also, due to the nature of the data, it was not possible to make conclusive arguments. Despite implemented measures, some participants might have struggled to differentiate subtle segmental differences in word pairs due to potential external auditory distractions in the test environment and insufficient training to complete the tasks provided. Future research could explore variabilities of learners including varied L1 backgrounds, language education backgrounds, and proficiency levels to provide a more comprehensive understanding.

Note

This article is based in part on findings from an unpublished presentation, titled “A *diagnostic assessment of pronunciation: Segmental and suprasegmental perceptions of EFL learners*”, which was delivered at the 15th METU International ELT Convention: Envisioning Future Paths in Ankara, Türkiye.

References

- Alnafisah, M., Goodale, E., Rehman, I., Levis, J., & Kochem, T. (2022). The impact of functional load and cumulative errors on listeners' judgments of comprehensibility and accentedness. *System*, 110, 1-14. <https://doi.org/10.1016/j.system.2022.102906>
- Arıkan, A., & Yılmaz, A. F. (2020). Pre-service English language teachers' problematic sounds. *International e-Journal of Educational Studies (IEJES)*, 4(7), 1-26. <https://doi.org/10.31458/iejes.594715>
- Avery, P., & Erlich, S. (1992). *Teaching American English pronunciation*. Oxford University Press.
- Baker, A. (2006). *Ship or sheep? An intermediate pronunciation course*. Cambridge University Press.
- Bardakçı, M. (2015). Turkish EFL pre-service teachers' pronunciation problems. *Educational Research and Reviews*, 10(16), 2370-2378. <https://doi.org/10.5897/ERR2015.2396>
- Barlow, J. A., & Gierut, J. A. (2002). Minimal pair approaches to phonological remediation. *Seminars in Speech and Language*, 23(1), 57-67. <https://doi.org/10.1055/s-2002-24969>
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14(3), 345-360. <https://doi.org/10.1037/0096-1523.14.3.345>
- Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, 61(5), 977-985.
- Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly* 22(4), 593-606. <https://doi.org/10.2307/3587258>
- Carr, P. (2008). *A glossary of phonology*. Edinburgh University Press.
- Catford, J. C. (1987). Phonetics and the teaching of pronunciation: a systemic description of English phonology. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 87-100). TESOL.
- Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2017). *Teaching pronunciation: A coursebook and a reference guide* (2nd Ed.). Cambridge University Press.
- Demirezen, M. (2005). Rehabilitating a fossilized pronunciation error: The /v/ and /w/ contrast by using the audio articulation method in teacher training in Turkey. *Journal of Language and Linguistic Studies*, 1(2), 183-192.
- Demirezen, M. (2008). The /æ/ and /ʌ/ phonemes as fossilized pronunciation errors for Turkish English language teachers and students: Undoing the fossilized pronunciation error. *Journal of Language and Linguistic Studies*, 4(2), 73-82.
- Demirezen, M. (2010). The causes of the schwa phoneme as a fossilized pronunciation problem for Turks. *Procedia Social and Behavioral Sciences*, 2, 1567-1571. <https://doi.org/10.1016/j.sbspro.2010.03.237>
- Derwing, T. M. (2023). An overview of pronunciation teaching and training. In U. K. Alves & J. I. A. Albuquerque (Eds.), *Second language pronunciation: Different approaches to teaching and training* (pp. 399-412). De Gruyter Mouton.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins Publishing Company.
- Ergenç, İ. (2002). *Konuşma dili ve Türkçenin söyleyiş sözlüğü*. Multilingual.
- Ergenç, İ., & Bekar Uzun, İ. (2017). *Türkçenin ses dizgesi: Ses, ünlüler, ünsüzler, konuşma organları*. Seçkin.
- Flege, J. E. (1988). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human communication and its disorders: A review* (pp. 224-401). York Press.

- Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233-277). York Press.
- Flege, J. E. (2003). A method for assessing the perception of vowels in a second language. In E. Fava & A. Mioni (Eds.), *Issues in clinical linguistics* (pp. 19-43). Unpress.
- Gökgöz-Kurt, B. (2023). The perception of the English bilabial glide phoneme /w/ among Turkish learners of L2 English: A preliminary study on success and contributing factors. *International Journal of Language Academy*, 11(4), 208-224. <https://doi.org/10.29228/ijla.72323>
- Haghigi, M., & Rahimy, R. (2017). The effect of L2 minimal pairs practice on Iranian intermediate EFL learners' pronunciation accuracy. *International Journal of Research in English Education*, 2(1), 42-48. <https://doi.org/10.18869/acadpub.ijree.2.1.42>
- Isbell, D. R. (2020). Diagnostic language assessment for L2 pronunciation: A worked example. In O. Kang, S. Staples, K. Yaw, & K. Hirschi (Eds.), *Proceedings of the 11th pronunciation in second language learning and teaching conference*, (pp. 127-140). Iowa State University
- Kang, O., & Kermad, A. (2018). Assessment in second language pronunciation. In O. Kang, R. I. Thomson, J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 511-526). Routledge.
- Kelly, G. (2012). *How to teach pronunciation* (14th Ed.). Pearson Longman.
- King, R. D. (1967). Functional load and sound change. *Language*, 43(4), 831-852. <https://doi.org/10.2307/411969>
- Kissling, E. M. (2014). What predicts the effectiveness of foreign-language pronunciation instruction? Investigating the role of perception and other individual differences. *The Canadian Modern Language Review*, 70(4), 532-558. <https://doi.org/10.3138/cmlr.2161>
- Knoch, U. (2017). What can pronunciation researchers learn from research into second language writing? In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 54-71). Multilingual Matters.
- Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.
- Levis, J. M., & Cortes, V. (2008). Minimal pairs in spoken corpora: Implications for pronunciation assessment and teaching. In C. A. Chapelle, Y. R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 1197-208). Iowa State University.
- Levis, J. M., & Barriuso, T. A. (2012). Nonnative speakers' pronunciation errors in spoken and read English. In J. M. Levis & K. LeVelle (Eds.), *Social factors in pronunciation acquisition, Proceedings of the 3rd pronunciation in second language learning and teaching conference*, (pp. 187-194), Iowa State University.
- Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, 34, 520-531. <https://doi.org/10.1016/j.system.2006.09.004>
- Przedlacka, J. (2018). An overview of phonetics for language teachers. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 39-57). Routledge.
- Putra, T. K., & Rochsantiningsih, D. (2018). Improving students' pronunciation ability in fricative sounds through minimal pair drill. *English Education Journal*, 7(1), 26-36.
- Roach, P. (1992). *Introducing phonetics*. Penguin English.
- Rogerson-Revell, P. (2018). English vowels and consonants. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 93-121). Routledge.
- Suzukida, Y., & Saito, K. (2021). Which segmental features matter for successful L2 comprehensibility?

- Revisiting and generalizing the pedagogical value of the functional load principle. *Language Teaching Research*, 25(3) 431-450. <https://doi.org/10.1177/1362168819858246>
- The Jamovi Project (2022). *Jamovi*. (Version 2.3) [Computer Software]. <https://www.jamovi.org>
- Trask, R. L. (1996). *A dictionary of phonetics and phonology*. Routledge.
- Uzun, T. (2022). The salient pronunciation errors and intelligibility of Turkish speakers in English. *MEXTESOL Journal*, 46(1), 1-15. <https://doi.org/10.61871/mj.v46n1-9>
- Uzun, T. (2024). Perception of prosodic speech features: final intonation and word stress for EFL learners. *Uluslararası Dil, Edebiyat ve Kültür Araştırmaları Dergisi (UDEKAD)*, 7 (2), 212-223. <https://doi.org/10.37999/udekad.1449612>
- Valenzuela, M. G., & French, P. (2023). Production of English vowel contrasts in Spanish L1 learners: A longitudinal study. *Loquens*, 10(1-2), 1-12. <https://doi.org/10.3989/loquens.2023.v10.i1-2>