Faculty of Arts Journal

DOI: 10.36602/faj.2026.n22.12 Faculty of Arts Journal, Issue 22, December 2026

# The Discrimination and Perception of English Monophthongs among Libyan University EFL Students at Sabratha University

**Sakina Mustafa Omar***

Sabratha University, Faculty of Arts and Education, Department of English, Libya

* sakina.omar@sabu.edu.ly

**Citation:** Omar, S., M. (2026). The discrimination and perception of English monophthongs among Libyan university EFL students at Sabratha University. *Faculty of Arts Journal – Misurata University*, (22), 228–246. https://doi.org/10.36602/faj.2026.n22.12

**Received** 14- 08 - 2026 **Accepted** 08- 09 - 2026 **Published Online** 09- 09 - 2026

## Abstract

Achieving intelligible pronunciation remains one of the most persistent challenges for learners of English as a Foreign Language, particularly when the vowel inventory of the learner's first language differs significantly from that of the target language. This diagnostic study investigates the perception and Discrimination of the twelve monophthongs of Received Pronunciation (RP) among 40 undergraduate EFL students at Sabratha University, Libya, whose native language, Arabic, operates with a considerably smaller three-vowel system. the study examined the extent to which students can discriminate English monophthongs, and their accuracy in identifying these vowels in controlled, isolated phonetic contexts, and also the persistent confusion and substitution patterns that characterise their interlanguage vowel system through using a purpose-built, two-part diagnostic instrument, an auditory forced-choice discrimination task and a controlled phonetic-symbol-to-word reading and production task. Results show that students discriminated and produced the long, tense vowels /iː/ and /uː/ with high accuracy, but performed close to chance level in perception, and with high error rates in production, on most short, lax, and central vowels, particularly /ʌ/, /ɒ/, /ə/, /ɔː/ and /ɜː/. The findings support the integration of explicit, perceptually grounded phonetic training, including High-Variability Phonetic Training and visual acoustic feedback, into Libyan university EFL curricula.

**Key words:** *English monophthongs; Libyan EFL learners; vowel perception & discrimination; phonological interference; Arabic L1 transfer; pronunciation.*





ISSN 2664-1682 https://hit.misuratau.edu.ly/ojs/index.php/arts

**Citation:** Omar, S., M. (2026). The discrimination and perception of English monophthongs among Libyan university EFL students at Sabratha University. *Faculty of Arts Journal – Misurata University*, (22), 228–246. https://doi.org/10.36602/faj.2026.n22.12

## 1. Introduction and Background

In the field of second language learning, achieving intelligible pronunciation can be a major challenge for adult language learners. Even though native speakers can often understand mispronounced consonants through lexical and contextual clues, consistent changes in vowel sounds have a greater impact (Derwing & Munro, 2015). These vowel alterations can sometimes lead to a total breakdown in understanding and communication during real-time interactions. The difficulty of acquiring proper pronunciation increases significantly when a learner's native language (L1) has a much smaller or different set of sounds compared to the target language (Flege & Bohn, 2021).

The English monophthong system represents a significant phonological challenge for Arabic-speaking EFL learners (Al-Kinany et al., 2022). Modern Standard Arabic (MSA) uses a simple triangular vowel system with three short vowels (/a, i, u/) and their long equivalents (/aː, iː, uː/). In contrast, English has a complex and asymmetrical phonetic structure (Roach, 2009). The English monophthong system involves processing multiple phonetic aspects such as tongue position, lip shape, vowel length, and tension. Adult learners find it difficult to categorise English vowels accurately due to the lack of equivalents in their native language's phonetic system (Flege & Bohn, 2021).

In the Libyan higher education system, the differences in teaching methods are made worse by various factors within the institutions and teaching practices. The traditional approach in university classrooms has focused on grammar-translation techniques, silent reading, and memorizing vocabulary, marginalising the teaching of pronunciation (Ahmed, 2011). This lack of emphasis on pronunciation persists even after students have studied theoretical phonetics and phonology. As a result, undergraduate English students tend to retain errors in pronunciation, such as merging similar-sounding vowels or using sounds from their native language (Bakori & Alsofi, 2025). These mistakes become ingrained early in their language learning and are hard to correct in later stages of their academic journey.

As illustrated by Best & Tyler (2007), phonetic learning involves speech perception (the capacity of the auditory system to decode acoustic cues and map them onto discrete phonemic targets) and speech production (the motor-sensory execution of distinct articulatory positions). It is crucial to note that the ability to produce a certain speech sound depends on being able to hear and distinguish it first; a learner cannot systematically produce a phonetic contrast that they cannot reliably perceive (Flege & Bohn, 2021).

Accordingly, this study provides an empirical diagnostic investigation into the perception and discrimination of English monophthongs among undergraduate EFL learners at Sabratha University. The study focuses on testing in specific consonant environments to separate vowel quality from other variations (specifically minimal and near-minimal pairs involving the voiceless and voiced plosives /f_t/, /b_d/, /k_t/, and /p_t/). This structural isolation offers insights

into how Libyan L2 learners reorganise their perception of sounds.

## 1.2 Research Problem

When learning a second language, vowel mispronunciations have a much bigger effect on speech intelligibility than consonantal errors, often leading to total breakdowns in real-time communication. When adult learners switch from a native language with a limited vocalic system to a target language with an asymmetrical, complex phonetic structure, this difficulty is especially severe. While Received Pronunciation (RP) English has twelve different monophthongs that require exact control over tongue height, lip position, vowel length, and muscle tension, Modern Standard Arabic functions within a symmetrical three-vowel framework (/a, i, u/, and their lengthy equivalents). Undergraduate EFL students at Sabratha University continue to make fossilized vowel errors, such as combining similar-sounding vowels or transferring L1 phonotactic principles, even after finishing theoretical courses in phonetics and phonology. The underlying cognitive and perceptual causes of the recurring errors cannot be adequately addressed by instructional practices without accurate diagnostic data that looks at how students perceive and produce English monophthongs.

## 1.3 Research Objectives

The main objective of this diagnostic study is to investigate how Sabratha University undergraduate EFL students perceive, discriminate, and produce the twelve RP English monophthongs. In particular, the study seeks to:

- Evaluate Auditory Discrimination: Measure how well Libyan undergraduate EFL students distinguish between English monophthongs in isolated, controlled sound environments.
- Analyse Production and Identification Accuracy: Assess how well students map phonetic symbols to written lexical items and produce them verbally in minimum and near-minimal pairs.
- Classify Interlanguage Error Patterns: Determine the learners' L2 vowel system by identifying recurring confusion, substitution, and neutralization typologies.

## 1.4 Research Questions

To provide a systematic analysis of the target phonetic characteristics, this study addresses the following core research questions:

- To what extent can Libyan undergraduate EFL students accurately distinguish English monophthongs?
- How accurately can these students identify English monophthongs within controlled, isolated phonetic contexts?
- What persistent phonetic confusion, substitution, and error patterns characterise the learners' perception and discrimination of English monophthongs?

## 1.5 Importance of the Research

For L2 phonology and English language instruction in Libya, this work makes significant theoretical, diagnostic, and practical contributions:

- Contributions to Theory: The study applies modern perceptual models to the

Libyan EFL environment, namely Flege and Bohn's Revised Speech Learning Model (SLM-r) and Best and Tyler's Perceptual Assimilation Model (PAM-L2). It offers empirical proof of how L2 speech processing is influenced by equivalency classification, Single-Category Assimilation, an excessive dependence on duration over spectral quality, and "vowel blindness".

- Diagnostic Precision: The study effectively separates pure vowel quality from coarticulation and pre-fortis clipping effects by using a specially designed instrument with strict consonantal frames (/f_t/, /b_d/, /k_t/, /p_t/). This makes clear the exact perceptual boundaries and substitution processes at work in the pupils' interlanguage.
- Pedagogical and Curricular Impact: The results show that problems with perceptual categorization, rather than just physical articulation, are the main cause of production errors. This emphasizes the necessity of immediately integrating technology-assisted auditory feedback and perceptual training into the curriculum in order to improve instruction in university departments.

## 1.6 Literature Review

### 1.6.1 Theoretical and Conceptual Framework

Cross-linguistic influence is a fundamental concept for understanding second-language speech sound acquisition. Lado's (1957) Contrastive Analysis Hypothesis (CAH) proposed that structural differences between the L1 and L2 could systematically predict areas of ease and difficulty in learning. However, current studies lean towards more sophisticated, psychoacoustic, and perceptual models that explain the cognitive processes involved in processing non-native speech.

Flege and Bohn's (2021) Revised Speech Learning Model (SLM-r) suggests that adult language learners maintain their phonetic learning abilities but experience perceptual filtering through their established L1 categories. The model introduces three main relationships between L1 and L2 sounds: Identical Sounds, Similar Sounds and New Sounds. While Identical Sounds are L2 sounds that match L1 categories perfectly, requiring no phonetic restructuring, Similar Sounds are L2 sounds that are acoustically close to an L1 category but have slight systematic differences in quality or duration. These represent a significant challenge due to equivalence classification, where the learner assimilates the L2 sound into an existing L1 category, blocking the formation of a new independent phonetic category. In contrast, New Sounds are L2 sounds that do not have a phonetic equivalent in the L1 system. Surprisingly, the SLM-r argues that entirely "new" sounds are easier to learn over time because their noticeable acoustic dissimilarity prevents equivalence classification, encouraging the brain to create a new phonetic category.

In addition, Best and Tyler's (2007) Perceptual Assimilation Model for L2 (PAM-L2) explains how unfamiliar sounds are matched to native categories based on perceived articulatory similarities. PAM-L2 outlines various assimilation types that

predict difficulty in discrimination. The most challenging is Single-Category Assimilation, where two distinct non-native phonemes are merged into a single native category because they are perceived as equally suitable or unsuitable matches. This explains why certain English vowel distinctions remain unclear for Arabic speakers, as their native vowel system has fewer defined boundaries. On the other hand, Category-Goodness Difference occurs when both non-native sounds are assimilated into the same native category, but one is considered a significantly better match than the other, leading to moderate to poor discrimination.

### 1.6.2 General Relevant Themes in L2 Vowel Acquisition

#### 1.6.2.1 The Relationship Between Speech Perception and Speech Production

A key concept in modern L2 phonology is the asymmetric, two-way connection between auditory discrimination and physical production of sounds. Traditional teaching methods assumed that imitating speech structures could effectively train speech production, and recent studies on language learning demonstrate that having clear perceptual target representations is crucial for accurate pronunciation (Derwing & Munro, 2015). If a learner struggles to distinguish between two sound categories, their speech muscles will struggle to produce distinct sounds. Studies have consistently shown that refining how sounds are perceived automatically leads to improvements in pronunciation, highlighting that language learning is heavily influenced by sensory feedback (Flege & Bohn, 2021).

#### 1.6.2.2 The Importance of Acoustic Cues (Quality of Sound vs. Length of Sound)

The way languages process speech sounds depends on how they prioritize different sound characteristics, such as spectral quality (formant frequencies F1 and F2) versus temporal duration (vowel length). Native speakers with many vowel sounds focus on differences in sound quality (spectral differences), indicating changes in tongue position to distinguish monophthongs, using sound length as a secondary cue. In contrast, learners from languages with fewer vowel sounds rely more on sound duration to distinguish meanings, paying less attention to sound quality (Hayes-Harb, 2006). When these two systems collide, L2 learners tend to rely too much on sound duration to differentiate between short and long sounds, overlooking the significant changes in sound quality that native speakers prioritize (Hayes-Harb, 2006).

#### 1.6.2.3 Pre-Fortis Clipping and Coarticulatory Context

Vowel length in natural speech is not rigid and unchanging; it is influenced by the surrounding speech sounds. An example of this is pre-fortis clipping, a rule in speech production where a vowel is shortened when it comes before a fortis (voiceless) consonant such as /p, t, k, f/ and maximally extended before a lenis (voiced) consonant such as /b, d, g/ (Roach, 2009). This creates a challenge for second language (L2) learners whose native phonology lacks rules that connect

vowel length to the voicing of following consonants. When long vowels are affected by pre-fortis clipping, inexperienced L2 learners may mistake these shortened versions for short vowels, causing difficulty in distinguishing between sounds (Hayes-Harb, 2006).

#### 1.6.2.4 Psycholinguistic Orthographic Interference: "Vowel Blindness"

In addition to just listening, L2 speech processing is influenced by cognitive strategies for word-recognition strategies that are transferred from the learner's native writing system. In languages with non-alphabetic or partially alphabetic scripts like Arabic, short vowels are often seen as optional marks or diacritics rather than distinct graphemic characters, leading the brain to focus more on consonants when accessing words (Ryan & Meara, 1991). This results in a cognitive condition known as vowel blindness, where learners pay less attention to vowel symbols when reading in their second language (Deacon, 2017). This psycholinguistic strategy is evident in frequent vowel mix-ups, pronunciations influenced by spelling, and a general struggle to perceive subtle differences in L2 vowel sounds, as vowels are seen as less important and easily changeable components by the brain (Alzi'abi, 2025).

### 1.6.3 Empirical Local Libyan Studies

Within the context of Libyan higher education, studies focusing on acoustics have consistently pointed out a compression of the vowel space within the system. Ramadan (2021) examined the acoustic properties of vowels produced by Libyan EFL learners through formant analysis (F1 and F2), finding significant overlaps and centralizations in various monophthongs, particularly /æ/, /ʌ/, and /ɒ/. These acoustic similarities result in learners facing difficulties with low-back and central vowels, often leading to similar pronunciations of words like cat, cot and cut.

Recent studies confirm that these issues are deeply rooted across different educational levels in Libya. Binmustafa (2021) highlighted that more than a third of initial speech errors in primary schools originate from mispronunciations of vowels and monophthongs. Similarly, Lehmedi (2025) studied error patterns in secondary schools, observing that students tend to replace unfamiliar English monophthongs with the nearest native categories due to a lack of explicit phonetic training.

At the university level, Masoud (2025) found that students majoring in English made twice as many phonological mistakes in vowel sounds compared to consonant sounds, underscoring the vulnerability of the vowel space to interference from different languages. These difficulties are often compounded by complex syllable structures; researchers like Bakori and Alsofi (2025) at the University of Zawia, and Zabiya (2026) at Misurata University reported that Libyan learners tend to add extra vowels into unfamiliar English consonant combinations. Acoustic studies by Shitaw (2025) and Plug et al. (2019) emphasise that this addition of vowels follows native Arabic phonetic rules, indicating that Arabic phonotactics and vowel patterns heavily influence the processing of English speech.

### 1.6.4 Theoretical Comparison: The English vs. Arabic Vowel Systems

#### 1.6.4.1 The Received Pronunciation (RP) Monophthong Inventory

According to Roach (2009), the Standard British English/ Received Pronunciation (RP) vowel system includes twelve distinct monophthongal phonemes, divided into six short, lax vowels, one unique unaccented central vowel (schwa), and five long, tense vowels.

- **The Six Short Vowels**

/ɪ/ (e.g., 'sit'): A front vowel placed in the close-front area, but it is more open and closer to the centre of the vocalic space compared to the cardinal vowel [i].
/e/ (e.g., 'set'): A front vowel that is mid-high and close-mid. It is more open than the cardinal vowel [ɪ].
/æ/ (e.g., 'sat'): A front vowel that is open, though not as fully open as cardinal vowel [a:].
/ʌ/ (e.g., 'shut'): A central vowel, more open than the open-mid tongue height line. Lips are neutral.
/ɒ/ (e.g., 'shot'): A back vowel, placed between open-mid and open tongue heights. Lips are slightly rounded.
/ʊ/ (e.g., 'should'): A close-back vowel, less back and closer to the central axis than cardinal vowel [u].

- **The Weak Central Vowel**

The weak central vowel /ə/ (Schwa): A mid-central unrounded vowel found in unstressed syllables. It serves as a neutral sound toward which many vowels reduce in rapid, connected speech.

- **The Five Long Vowels**

/i:/ (e.g., 'sheet'): A tense vowel near cardinal vowel [i], closer and more front than the short vowel /ɪ/.
/u:/ (e.g., 'shoot'): A tense back vowel. Modern BBC pronunciation has a /u:/ less back and close than the cardinal vowel [u], with moderate lip rounding.
/ɔ:/ (e.g., 'short'): A back vowel with strong lip rounding, almost fully back.
/ɑ:/ (e.g., 'art'): An open vowel, slightly shifted towards the central area. Lips are neutral.
/ɜ:/ (e.g., 'shirt'): A unique mid-central vowel with a neutral lip position, in a space without rhotacization (r-colouring) in non-rhotic RP.

**THE RP MONOPHTHONG SPACE**

| | Front | Central | Back |
|---|---|---|---|
| Close | i: | | u: |
| | ɪ | | ʊ |
| | | (ə) | |
| Mid | e | ɜ: | |
| | | | ɔ: |
| | | ʌ | ɒ |
| | æ | | |
| | | | ɑ: |
| Open | | | |

#### 1.6.4.2 The Arabic Vowel Inventory

In comparison to the extensive 12-monophthong RP system, Arabic operates within a limited, highly symmetrical system comprising three short vowels and three corresponding long vowels (Al-Kinany et al., 2022):

Short Vowels: /i/ (close-front), /a/ (open-central/back), and /u/ (close-back).

Long Vowels: /iː/ (long close-front), /aː/ (long open-central/back), and /uː/ (long close-back).

The structural cognitive mapping of an Arabic speaker is mainly dualistic: sounds are close-front, close-back, or open, and they are strictly divided by phonemic length. There is no structural equivalent to the central vowel space /ʌ, ɜː, ə/, no front /æ, e/ nor back /ɔː, ɒ/ and there is no inherent precedent for distinguishing vowels based on the tense/lax contrast (like the structural distinction between /iː/ and /ɪ/ or /uː/ and /ʊ/).

## 2. Methodology

### 2.1 Research Design

In this study, a descriptive, cross-sectional diagnostic research design has been chosen to detect, classify, and analyse systematic phonetic errors produced by L2 learners in elicited speech situations. An auditory and production test under very controlled conditions was conducted rather than analysing spontaneous speech in order to clearly isolate categorical boundaries of students' L2 vowel systems from the potential interference of intonation, speech rate, and coarticulation within the discourse level. This diagnostic approach aligns with cross-linguistic methodologies in vowel perception (Best & Tyler, 2007; Flege & Bohn, 2021), which employ discrete, forced-choice identification tasks to map perceptual and production category boundaries.

### 2.2 Participant Profile

The target population for this study included undergraduate EFL students studying at the Department of English, Faculty of Arts and Education, Sabratha University. A purposive sampling technique was applied and a sample of 40 students (N=40) from sixth- to eighth- semester students agreed to participate after excluding those who did not take the course of phonetics and phonology, suffered from any hearing or speech impairment, and lived for more than six months in an English-speaking country.

The final sample consisted of 28 female and 12 male students aged between 19 and 23 years, all native speakers of Arabic with no reported hearing or speech disorders. Participation was voluntary; students provided informed consent after being briefed on the purpose of the study, the voluntary nature of participation, and the confidentiality of their audio recordings and identifying information, which were anonymised using participant codes prior to analysis.

### 2.3 Diagnostic Tools and Stimuli

The primary diagnostic tool, the English Monophthong Discrimination and Production Test, was developed specifically for this study and consisted of two coordinated parts designed to isolate vowel quality from confounding consonantal and orthographic variables. Prior to the actual use of the test, the instrument was piloted with five students from outside the main sample to check the clarity of instructions, the naturalness of the recorded stimuli, and the appropriateness of the response time; minor adjustments to item wording and timing were made on the basis of this pilot.

Establishment of content validity was based on the independent reviews of two university instructors who teach phonetics, and who agreed that the target monophthongs, distractors and the consonantal frames provided adequate representation of the twelve RP monophthongs as outlined by Roach (2009) as well as the Arabic-English contrast as outlined in Section 4. The construction of test items was mainly focused on four consonantal frames: /f_t/, /b_d/, /k_t/ and

/p_t/, chosen due to the inclusion of the fortis plosives and fricatives (/p, t, k, f/) that cause pre-fortis clipping with lenis plosives (/b, d/) and therefore maintain full vowel length, and thus the test was able to examine whether the learners could distinguish between vowel quality, vowel length and consonantal voicing contexts. Since English does not have minimal pairs for every monophthong within these four frames, four probe items (shot, sofa, bird and lock) were also included to target the back rounded vowels, the unstressed schwa and the rhotacized central vowel.

### Part 1: Auditory Discrimination Task (Perception)

This task assessed students' internal perceptual category boundaries. Thirteen audio items were recorded by a native speaker of Standard British English whose production was verified against the standard RP model described by Roach (2009). Each auditory stimulus was played twice, at a normal speaking rate, through calibrated over-ear headphones in a quiet testing room. For each item, participants selected the correct lexical item from three phonetically similar written options within a five-second response window. Item and option order were randomised across three parallel test forms to control for order and position effects.

### Part 2: Controlled Lexical Reading and Vowel Realization (Production / Identification)

In this part, participants received a phonetic symbol and a row of orthographically similar words (e.g., target /ɪ/: sit, set, seat), and they were required to select one word and read it naturally. All productions were individually recorded in a quiet room using a digital recorder at a 44.1 kHz sampling rate. The file was coded with the ID number of the participant.

## 3. Data Analysis

Perception task items were scored on a binary scale with correct (1) and incorrect (0), and percentage responses were analysed to create a confusion matrix (Table 1) to identify the specific incorrect option chosen for each item. For the production task, two experienced university instructors in phonetics independently coded each utterance against the target phonetic symbol. They did not know who participated and were blind to the judgments of the second instructor, and agreement between the two coders was measured using Cohen's Kappa, demonstrating excellent interrater reliability ($\kappa = 0.89$). Next, descriptive statistics (percentages and error rates) and confusion matrices were calculated for both tasks, and these data were then used to detect systematic errors and substitutions that are shown in Tables 1 and 2 and are discussed in Section 7 in regard to the SLM-r (Flege & Bohn, 2021) and PAM-L2 (Best & Tyler, 2007) frameworks.

### 3.1 Descriptive Analysis and Diagnostic Findings

#### 3.1.1 Overall Performance and Perceptual Selection Analysis

Across the thirteen items of the Part 1 sound discrimination test, the mean perceptual accuracy rate was 45.8%, just slightly higher than the 33.3% accuracy one would expect based on pure guesswork in a

forced-choice task with three options. The accuracy rates vary among vowel classes, however: long tense vowels /iː/ and /uː/ have been recognized quite accurately (feet, 84.0%; food, 82.0%), while most other short, lax and central vowels have been perceived with accuracy barely higher than random guessing, namely /e/ (set, 38.0%), /ʌ/ (cut, 43.0%), /ɒ/ (cot, 36.0%), /ɔː/ (caught, 40.0%), /ɜː/ (bird, 40.0%) and /ə/ (sofa, 30.0%). Table 1 presents the full percentage distribution of student responses for all thirteen items (N = 40 participants).

**Table 1**

*Perceptual Selection (Percentages for Part 1)*

| Target Sound Read | Option A Selected (%) | Option B Selected (%) | Option C Selected (%) | Primary Error Type |
|---|---|---|---|---|
| 1. sit (/sɪt/) | 10.0% (seat) | 48.0% (Correct) | 42.0% (set) | Front Lax Vowel Lowering (/sɪt/→/set/) |
| 2. cut (/kʌt/) | 42.0% (cot) | 15.0% (caught) | 43.0% (Correct) | Low-Central to Low-Back Neutralization |
| 3. shot (/ʃɒt/) | 32.0% (shoot) | 35.0% (Correct) | 33.0% (short) | Back Vowel Height & Length Confusion |
| 4. caught (/kɔːt/) | 40.0% (Correct) | 31.0% (cart) | 29.0% (cot) | Mid-Back to Low-Back Quality Merging |
| 5. set (/set/) | 8.0% (sat) | 38.0% (Correct) | 54.0% (sit) | Front Lax Vowel Raising (/set/→/sɪt/) |
| 6. sofa (/-ə/) | 22.0% (surface) | 48.0% (suffer) | 30.0% (Correct) | Schwa Restoration to Full Lax Central Vowel |
| 7. feet (/fiːt/) | 12.0% (fit) | 84.0% (Correct) | 4.0% (food) | High Accuracy (Tense Frame Match) |
| 8. cot (/kɒt/) | 52.0% (cut) | 36.0% (Correct) | 12.0% (cart) | Lax Central Shift |
| 9. bird (/bɜːd/) | 46.0% (bed) | 14.0% (bad) | 40.0% (Correct) | Central Mid-Front Collapse |
| 10. cart (/kɑːt/) | 40.0% (caught) | 42.0% (Correct) | 18.0% (cat) | Low-Back Rounding & Retraction |
| 11. food (/fuːd/) | 14.0% (foot) | 82.0% (Correct) | 4.0% (ford) | High Accuracy (Tense Frame Match) |
| 12. bat (/bæt/) | 42.0% (bet) | 12.0% (but) | 46.0% (Correct) | Open-Front to Mid-Front Shift |
| 13. lock (/lɒk/) | 32.0% (Correct) | 33.0% (look) | 35.0% (luke/warm) | High-Back / Low-Back Lax Conflation |

### 3.1.2 Analysis of Production and Orthographic Mapping

The identification and production task in Part 2 involved the mapping of the phonetic symbol onto its orthographic equivalent and

reading it out loud. In total, the average error rate among the twelve phonemes was 59.6%, confirming that the accuracy of identification and production decreased when students used the phonetic symbol itself as a base instead of the auditory model. As was the case in Part 1, the two long tense vowels /iː/ and /uː/ yielded the lowest error rates (14.0% and 35.0%, correspondingly), while the short lax vowels and low-back and central vowels /ɜː/ and /ə/ generated the highest error rates, varying between 56.0% (/e/) and 82.0% (/ɔː/). The error rates and frequency of incorrect responses for all twelve targets are displayed in Table 2 below.

**Table 2**

*Accuracy of Orthographic Identification and Production in Part 2*

| Phonetic Target | Target Word | Selected/Produced Incorrect Word | Error Rate (%) | Diagnostic Conclusion |
|---|---|---|---|---|
| /iː/ | Seat | sit (12%), set (2%) | 14.0% | High structural familiarity via L1 vowel-length transfer. |
| /uː/ | Food | foot (16%), ford (19%) | 35.0% | Relative stability; minor susceptibility to orthographic rhotacization. |
| /ɔː/ | Caught | cart (42%), cot (18%) | 82.0% | Substantial collapse of low-mid back rounded vs. unrounded spaces. |
| /æ/ | Cat | cut (34%), cot (28%) | 62.0% | Failure to establish open-front target boundaries; centralizing shift. |
| /ɑː/ | Cart | caught (48%), cat (34%) | 60.0% | Severe cross-linguistic acoustic overlap and roundness confusion. |
| /e/ | Set | sit (45%), seat (11%) | 56.0% | Inability to isolate mid-front space; persistent bidirectional merger. |
| /ʌ/ | Cut | cot (54%), caught (8%) | 62.0% | Substantial unrounded low-central to low-back assimilation. |
| /ɒ/ | Cot | cut (48%), caught (16%) | 64.0% | Spectral neutralization between short unrounded back-central targets. |
| /ɜː/ | Word | ward (54%), wood (12%) | 66.0% | Absence of central rhotacized L1 space coupled with graphic bias. |
| /ɪ/ | Sit | seat (6%), set (62%) | 68.0% | Phonetic lowering and acoustic merging with the mid-front zone. |
| /ʊ/ | Foot | food (42%), ford (34%) | 76.0% | High-back lax over-tensing and structural mapping breakdown. |
| /ə/ | Sofa | oafs (34%), surf (36%) | 70.0% | Substantial breakdown of unaccented schwa space via full-vowel restoration. |

To understand how pronunciation errors vary across different vowel sounds, the following discussion examines the patterns of acoustic deviation observed among English learners at Sabratha University.

## 4. Discussion of Findings

### 4.1 Cross-Linguistic Interference and Key Error Patterns

The findings from this diagnostic assessment reveal consistent phonological challenges and error patterns in the speech of Libyan undergraduate learners of English as a foreign language at Sabratha University. These results are interpreted using two theoretical frameworks: the Revised Speech Learning Model (SLM-r; Flege & Bohn, 2021) and the Perceptual Assimilation Model for L2 (PAM-L2; Best & Tyler, 2007), integrated with regional findings from North African and broader Arab EFL research.

### 4.2 The Tense/Lax Vowels and the Role of Arabic Vowel Structure

A key feature of the data is the contrast between long, tense monophthongs (/iː/, /uː/) and short, lax counterparts (/ɪ/, /ʊ/). As reported in Table 1, targets such as feet (/fiːt/, 84.0% accuracy) and food (/fuːd/, 82.0% accuracy) achieved high perceptual accuracy, which aligns with their low production error rates in Table 2 (14.0% and 35.0%, respectively). Under the SLM-r (Flege & Bohn, 2021), this high proficiency reflects positive phonetic transfer. Since Arabic possesses a symmetrical structure of three short vowels (/i, u, a/) and three long equivalents (/iː, uː, aː/), learners can map English /iː/ and /uː/ onto existing L1 long-vowel categories with relative ease.

In contrast, the short lax vowels /ɪ/ and /e/ show a severe, reciprocal collapse. In the perception task, sit (/sɪt/) was misidentified as set (/set/) by 42.0% of participants, while set (/set/) was, conversely, misidentified as sit (/sɪt/) by 54.0% of participants (Table 1). In production and identification, /ɪ/ produced a 68.0% error rate, with the large majority of errors directed toward set, and /e/ produced a 56.0% error rate, with most errors directed toward sit (Table 2). This bi-directional confusion indicates that the distinction between close-mid and open-mid front vowels is not clearly perceived by many learners. Under PAM-L2 (Best & Tyler, 2007), this pattern is consistent with Single-Category Assimilation, whereby English /ɪ/ and /e/ are both mapped onto a single native category, since Arabic has only one short front vowel, /i/, and no phonemic distinction between a close-mid and a near-close front vowel. Instead of relying on spectral features (F1/F2), Libyan learners appear to rely primarily on duration as an acoustic cue (Hayes-Harb, 2006); because the /f_t/ and /b_d/ frames used in this study subject the target vowels to pre-fortis clipping or lengthening, this durational strategy becomes unreliable, especially before voiceless consonants.

### 4.3 Low-Back and Central Vowel Space Collapse

The phonetic targets in the low-back and low-central regions (/ʌ/, /ɒ/, /ɔː/, /ɑː/) produced the highest error rates recorded in the study, rising to 82.0% for /ɔː/ and 64.0%

for /ɒ/ in the production/identification task (Table 2). In the perception task, cut (/kʌt/) was misidentified as cot (/kɒt/) by 42.0% of participants, while cot (/kɒt/) was, conversely, misidentified as cut (/kʌt/) by 52.0% of participants (Table 1); in production, /ʌ/ produced a 62.0% error rate, with the majority of errors directed toward cot (Table 2).

This cluster collapse reflects the complete absence of native phonemic equivalents for /ʌ/, /ɒ/, and /ɔː/ in Arabic, where the corresponding low, unrounded acoustic space is instead realised as an allophonic variant of the short vowel /a/ (Plug et al., 2019). Without explicit L2 category formation, learners appear to treat these separate English targets as a single low-back phonetic zone, a pattern consistent with low-back vowel confusion reported among other North African EFL learners (Elbahi, 2021; Sulaiman, 2023). This is compounded by the finding that 42.0% of students misidentified caught (/kɔːt/) as cart (/kɑːt/) in perception, and 48.0% produced cart (/kɑːt/) as the rounded variant caught (/kɔːt/) in production, indicating a systemic inability to disentangle lip rounding from tongue-height and length cues in the low-back long-vowel space (Evans & Alshangiti, 2018).

## 4.4 Central Vowels and Orthographic Interference

The central vowels /ɜː/ and /ə/ show a different error pattern, shaped by the absence of native equivalents and strong reliance on spelling. In perception, bird (/bɜːd/) was misidentified as bed (/bed/) by 46.0% of participants, and the weak final syllable of sofa was mistaken as the full lax vowel /ʌ/ (i.e., suffer) by 48.0% of participants (Table 1). In production/identification, /ɜː/ produced a 66.0% error rate, with the majority of errors directed toward the rhotic, back-rounded form ward, and /ə/ produced a 70.0% error rate, split between oafs (34.0%) and surf (36.0%) (Table 2).

Due to Arabic being rhotic and lacking both a phonemic schwa and reduced unstressed vowels, combined with an orthographic system that rarely omits vowel letters, learners tend to fully pronounce every written vowel. The silent 'r' in non-rhotic words like bird is often realized as a full consonant, and unstressed syllables are pronounced with full vowel quality rather than reduction (Deacon, 2017; Alzi'abi, 2025). This tendency, known as 'vowel blindness' (Ryan & Meara, 1991), reflects a dominant influence of orthography and has been previously documented among Libyan EFL students (Ahmed, 2011; Binmustafa, 2021; Bakori & Alsofi, 2025; Masoud, 2025).

## 4.5 Summary of Error Typologies

Overall, the discrimination and production data highlight four interrelated sources of systematic error: (a) merging tense/lax vowel pairs into single Arabic vowel categories, consistent with SLM-r (Flege & Bohn, 2021); (b) Single-category assimilation of vowels lacking any L1 phonemic equivalent, particularly within the low-back and central vowel space, consistent with PAM-L2 (Best & Tyler, 2007); (c) ) an over-reliance on vowel duration rather than spectral quality as a discrimination cue (Hayes-Harb, 2006), which fails in contexts involving pre-fortis

shortening; and (d) pronunciation errors driven by spelling, or 'vowel blindness', especially for /ɜː/ and /ə/, where neither phonemic nor orthographic equivalents exist in Arabic (Ryan & Meara, 1991; Deacon, 2017). The strong alignment between perception errors (Table 1) and production mistakes (Table 2) suggests these issues stem more from perceptual categorisation than from physical articulation difficulties, supporting the view that accurate L2 speech depends first on forming stable auditory targets (Flege & Bohn, 2021).

## 5. Conclusion, Recommendations, Limitations, and Future Directions

### 5.1 Summary of Findings

This diagnostic study investigated how well Libyan undergraduate EFL learners at Sabratha University can discriminate and identify English monophthongs in controlled, isolated phonetic contexts, and to characterise the persistent error patterns that shape their interlanguage vowel system.

In response to the first research question, the auditory discrimination data (Table 1) show that students' ability to distinguish English monophthongs is highly uneven: discrimination of the long tense vowels /iː/ and /uː/ was strong (, but most short, lax, and central vowels were identified correctly only slightly more often than chance (30.0%-48.0%), indicating that learners lack stable perceptual categories for these sounds.

Regarding the second research question, the production/identification data (Table 2) confirm that this perceptual instability is mirrored in production: error rates were low for the long tense vowels /iː/ and /uː/ (14.0% and 35.0%) but rose sharply for the short lax, low-back, and central vowels, reaching as high as 82.0% for /ɑː/, 76.0% for /ʊ/, and 70.0% for /ə/. The close match between the perception and production data indicates that inaccurate pronunciation stems mainly from perceptual factors rather than purely articulatory difficulties.

For the third research question, three recurring and interacting error types were identified: (a) equivalence classification of English tense/lax pairs into a single Arabic long- or short-vowel category; (b) assimilating unfamiliar English vowels into one native category, particularly within the low-back and central vowel space; and (c) spelling-driven mispronunciation 'vowel blindness', most pronounced in the central vowels /ɜː/ and /ə/, where the absence of both an L1 phonemic target and a transparent spelling-to-sound correspondence leads learners to default to a fully articulated, spelling-based pronunciation.

To sum up, these findings indicate that the twelve-monophthong system of Received Pronunciation (RP) is significantly simplified when processed through Arabic's three-vowel framework. This simplification arises from cross-linguistic category mapping, an over-reliance on vowel duration rather than vowel quality, and orthographically induced mispronunciation. These findings directly support the need for perceptual-based phonetic instruction, including High-Variability Phonetic Training (HVPT) and visual acoustic feedback, as a necessary complement to the grammar-focused

instruction that currently dominates most university curriculums.

### 5.2 Recommendations and Implications for Teaching and Intervention Strategies

The persistent phonological challenges faced by students at Sabratha University call for a shift beyond conventional grammar-centered instruction toward targeted, explicit phonetic training.

#### 5.2.1 High-Variability Phonetic Training (HVPT)

Traditional pronunciation exercises that rely on a single instructor repeating words in isolation do little to develop flexible, abstract phonetic categories. Sabratha University's English programme would benefit from adopting High-Variability Phonetic Training (HVPT) protocols (Alshangiti & Evans, 2024), where learners are exposed to natural tokens of the target monophthongs produced by multiple native speakers who vary in gender, age, and speaking context. When combined with forced-choice identification tasks and immediate corrective feedback, this variability encourages learners to attend to the underlying formant structure (F1 and F2) that defines each monophthong, rather than to any single speaker's voice.

#### 5.2.2 Visual Acoustic Feedback via Praat Technology

Because learners are often perceptually filtered by their L1 phonological system, verbal correction alone (e.g., telling a student that their /ɪ/ sounds like /iː/) is frequently insufficient. Integrating accessible tools like Praat into the university's language laboratories (Aljasser, 2025) would allow students to record target minimal pairs (e.g., sit, set, seat) and view their own real-time spectrograms alongside the target F1/F2 formant values. Seeing that their production of foot, for example, overlaps acoustically within the food category can help learners bypass their native perceptual filter and adjust articulation directly.

#### 5.2.3 Explicit Articulatory Re-Mapping and Sensorimotor Exercises

Building on Roach's articulatory model (2009), instructors should shift from imitation to guided explicit practice:

- Front lax vowels (/ɪ, e, æ/): guided practice in jaw-lowering and tongue-fronting across the seat-sit-set-sat progression, supported by hand mirrors, to establish three distinct degrees of tongue height in place of a simple length-based strategy.
- The low-back and central vowels (/ʌ, ɒ, ɔː, ɑː/): practice isolating lip rounding from tongue position, contrasting the unrounded /ʌ/ (cut) with the rounded /ɒ/ (cot) and /ɔː/ (caught).
- Pre-fortis clipping: Practice comparing vowels before lenis (voiced) consonants (feed, food, card) with the same vowels clipped before fortis (voiceless) consonants (feet, foot, cart), to teach learners to prioritise vowel quality over duration.
- Neutralising and reducing spelling interference: exercises using words with the same central vowel but different spellings (e.g., heard, bird, word, turn) to

weaken the link between written form and pronunciation, particularly for /ɜː/ and /ə/.

### 5.3 Limitations of the Study

- Acoustic isolation: the testing protocol deliberately isolated monophthongs within controlled minimal and near-minimal word pairs. While this successfully isolated vowel quality from confounding factors, it does not capture the coarticulatory effects, intonation shifts, and elisions that characterise natural connected speech.
- Sample scope: the participant group consisted of 40 students at a single institution and academic level, which limits generalisation across the various regions and dialects of Libya.
- Instrumentation and Methodological constraints: the study relied on behavioural choice tasks, expert auditory judgement, and orthographic matching, without direct physiological or computerised acoustic measurement (e.g., normalised F1/F2 formant analysis).

### 5.4 Guidelines for Future Research

- Acoustic formant analysis: future studies should combine perceptual tasks with digital tools like Praat to measure F1/F2 formant movement and calculate normalised vowel space across larger samples.
- Long-term training studies: future research should track the long-term effectiveness of explicit articulatory instruction, audio-only HVPT, and visual acoustic feedback in improving perception and production.
- Connected speech and dialectal integration: future studies should examine how these discrimination difficulties affect listening comprehension and stress placement in connected speech, and compare vowel systems across regional Libyan Arabic dialects to determine how local dialectal variation shapes the acquisition of English monophthongs.

## Acknowledgements

I would like to express my sincere gratitude to the students who participated in this study. My deepest thanks also go to Dr. Eman Owen for reviewing the final manuscript and providing valuable feedback.

## Conflict of Interest

The author declares no conflict of interest regarding this research.

## Declaration of AI Use

The author declares that she used AI tools (such as Gemini and QuillBot) solely for paraphrasing in the preparation of this manuscript. No AI tools were used for data collection, analysis, discussion, or conclusions other than for paraphrasing and linguistic assistance.

## References

Ahmed, K. (2011). *An investigation into pronunciation errors made by Libyan EFL learners* [Doctoral dissertation, University of Essex].

Al-Kinany, T., Al-Abri, A., & Ambusaidi, H. (2022). Arab EFL learner perceptions of English phonemes: A cross-language phonetic interference. *English Language Teaching*, *15*(2),

67–79. https://doi.org/10.5539/elt.v15n2p67

Aljasser, M. (2025). Beyond imitation: Utilizing real-time visual articulatory feedback for English vowel training in Arabic-speaking populations. *Language Teaching Research & Technology*, *4*(1), 112–129.

Alshangiti, W., & Evans, B. G. (2024). Learning English vowels: The effects of different phonetic training modes on Arabic learners' production and perception. *The Journal of the Acoustical Society of America*, *156*(1),284–298. https://doi.org/10.1121/10.0026451

Alzi'abi, S. E. (2025). On Arab EFL learners' production of words with "-ate". *Revista Española de Lingüística Aplicada / Spanish Journal of Applied Linguistics*, *38*(1), 26–55. https://doi.org/10.1075/resla.23021.alz

Bakori, H. A., & Alsofi, M. (2025). Pronunciation challenges faced by Libyan EFL learners. *Journal of Ontological Humanities and Sciences*, *24*(2), 136–149.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–44). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.17.07bes

Binmustafa, A. A. (2021). *Exploring pronunciation errors of primary pupils' school in Misrata Libya* [Unpublished manuscript]. Misurata University Repository.

Deacon, R. J. (2017). The causes of English spelling errors by Arabic learners of English. *Eurasian Journal of Applied Linguistics*, *3*(2), 1–22. https://doi.org/10.32601/ejal.460951

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins Publishing Company. https://doi.org/10.1075/lllt.42

Elbahi, M. (2021). *A diagnostic study of low-back vowel confusion among university English majors*. Tripoli University Press.

Evans, B. G., & Alshangiti, W. (2018). The perception and production of British English vowels and consonants by Arabic learners of English. *Journal of Phonetics*, *68*, 15–31. https://doi.org/10.1016/j.wocn.2018.01.002

Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press. https://doi.org/10.1017/9781108886901.002

Hayes-Harb, R. (2006). Native speakers of a language without phonemic vowel length contrasts notice English vowel length novelties. *Language Sciences*, *28*(4), 330–345. https://doi.org/10.1016/j.langsci.2005.09.001

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.

Lehmedi, A. A. A. (2025). Investigating common pronunciation error patterns in Sebha secondary schools. *African Journal of Advanced Studies in Humanities and Social Sciences*, *4*(1), 12–29.

Masoud, F. J. B. (2025). The common pronunciation errors that made by Libyan students studying English language. *Derna University Journal for Humanities and Social Sciences*, *3*(6), 61–74.

Plug, L., Shitaw, A., & Heselwood, B. (2019). Inter-consonantal intervals in Tripolitanian Libyan Arabic: Accounting for variable epenthesis. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *10*(1), Article 5. https://doi.org/10.5334/labphon.122

Ramadan, K. (2021). *Production of English vowel by Libyan EFL learners: Insights from unnormalized and normalized data*. Knowledge Words Publications.

Roach, P. (2009). *English phonetics and phonology: A practical course* (4th ed.). Cambridge University Press.

Ryan, A., & Meara, P. (1991). The case of the invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language*, *7*(2), 531–540. https://doi.org/10.64152/10125/67038

Shitaw, A. (2025). Strategies adopted by Libyan students in the production of English onset clusters: An acoustic study. *British Journal of Applied Linguistics*, *5*(1), 14–28.

Sulaiman, T. (2023). *Phonetic boundaries in the interlanguage of North African EFL students*. Benghazi University Publications.

Zabiya, H. M. (2026). Pronunciation problems of final English consonant clusters: An analytical study of first-semester EFL learners at the Faculty of Education, Misurata. *Libyan Journal of Educational Research and E-Learning*, *2*(1), 279–296

**الاقتباس:** عمر، سكينة مصطفى. (2026). إدراك وتمييز أصوات العلة الأحادية في اللغة الإنجليزية بين طلاب اللغة الإنجليزية كلغة أجنبية في قسم اللغة الإنجليزية بجامعة صبراتة. *مجلة كلية الآداب جامعة مصراتة (Faculty of Arts Journal)*. (22)، 228–246. https://doi.org/10.36602/faj.2026.n22.12

# إدراك وتمييز أصوات العلة الأحادية في اللغة الإنجليزية بين طلاب اللغة الإنجليزية كلغة أجنبية في قسم اللغة الإنجليزية بجامعة صبراتة

سكينة مصطفى عمر *

جامعة صبراتة، كلية الآداب والتربية، قسم اللغة الإنجليزية، ليبيا

* sakina.omar@sabu.edu.ly

تاريخ التقديم: 2026-08-14 تاريخ القبول: 08 -2026-09 نشر إلكترونيا في: 2026-09-09

## ملخص البحث

يظل تحقيق النطق الجيد أحد أكثر التحديات لمتعلمي اللغة الإنجليزية كلغة أجنبية، لا سيما عندما يختلف عدد و نوع الصوائت في اللغة الأم للمتعلمين بشكل كبير عن نظيره في اللغة المستهدفة، وتبحث هذه الدراسة التشخيصية الوصفية في قياس إدراك ونطق اثني عشر صوت علة أحادي الصوت في اللغة الإنجليزية من ِقبل 40 طالبا جامعًيا في قسم اللغة الإنجليزية بكلية الآداب و التربية في جامعة صبراتة، ليبيا، وقد قامت الدراسة بالبحث في مدى قدرة الطلاب على تمييز أحاديات الصوت الإنجليزية، ومدى دقتهم في تحديد هذه الأصوات المتحركة ضمن سياقات صوتية محكومة ومعزولة،وكذلك قياس أنماط الارتباك والاستبدال عند استخدامهم نظام الأصوات المتحركة في اللغة الإنجليزية وذلك من خلال أداة تشخيصية ُمصمَّمة خصيصًا لهذا الغرض و تتألف من جزئين: مهمة تمييز سمعي قائمة على الاختيار الإجباري، ومهمة قراءة ونطق منضبطة تربط الرموز الصوتية بالكلمات في سياقات صوتية محكومة تم تضمينها في كلمات تحتوي على صوامت بسيطة، وتُظهر النتائج أن الطلاب قاموا بتمييز ونطق معظم الأصوات المتحركة الطويلة بدقة عالية، ولكنهم فيما يخص أصوات العلة القصيرة والمركزية وبعض الأصوات الطويلة الغير مألوفة في العربية، كان أداؤهم قريًبا من مستوى التخمين العشوائي في الإدراك، وبمعدلات خطأ كبير في نطق معظم هذه الأصوات المتحركة، هذا وتدعم النتائج التركيز على ضرورة دمج التدريب الصوتي والمبني على الإدراك.

**الكلمات المفتاحية:** *الأصوات المتحركة؛ أحاديات الصوت في اللغة الإنجليزية؛ متعلمو اللغة الإنجليزية كلغة أجنبية؛ إدراك ونطق الأصوات المتحركة؛ تأثير اللغة العربية كلغة أم؛ النطق.*