Journal of English Language Teaching and Applied Linguistics
ISSN: 2707-756X
DOI: 10.32996/jeltal
Journal Homepage: www.al-kindipublisher.com/index.php/jeltal

JELTAL
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

A Systematic Review of Studies on Pronunciation Instruction and Practice in L2 (2005-2025)

Reima Al-Jarf
Full Professor of English and Translation Studies, Riyadh, Saudi Arabia
Corresponding Author: Reima Al-Jarf, **E-mail**: reima.al.jarf@gmail.com

| ABSTRACT

This systematic review (SR) synthesizes sixteen research articles published by the author between 2005 and 2025 on pronunciation pedagogy and practice in L2 contexts, with a particular focus on Arabic–English learners. The studies covers multiple instructional settings and modalities and were categorized into five thematic clusters: (i) AI generated speech and pronunciation errors, (ii) segmental errors in specialized domains, (iii) prosody, discourse, and spoken Arabic phenomena, (iv) technology enhanced pronunciation and listening practice, and (v) cognitive foundations of listening, decoding, and pronunciation. Across these clusters, the findings converge on a central issue: pronunciation is not an isolated phonetic skill but a cognitive–linguistic process shaped by phonological transfer, lexical access, decoding efficiency, background knowledge, and input quality. When any of these components is weak, learners rely on predictable compensatory strategies, such as nonsense forms, rhyming analogies, literal translation, or overgeneralization, revealing the cognitive pressures involved in real time speech production. The SR also shows that effective pronunciation pedagogy extends beyond segmental practice. Instruction that integrates decoding, grapheme–phoneme correspondences, sustained exposure to accurate models, and guided oral reading results in more consistent improvement in accuracy and fluency. Technology as text to speech software, YouTube videos, mobile apps and AI generated speech, emerges as a promising support for pronunciation practice when learners engage with it regularly and receive structured opportunities to connect written and spoken forms. Nevertheless, several gaps still exist in the field: suprasegmental features are underrepresented, longitudinal studies are limited, and many technology based interventions lack rigorous phonological evaluation. Addressing these gaps is essential for advancing pronunciation pedagogy. Overall, this study contributes a clearer, more integrated understanding of pronunciation pedagogy by positioning pronunciation as a cognitive linguistic skill shaped by perception, decoding, vocabulary development and repeated exposure to accurate models. The review offers a foundation for future research and provides educators with evidence based insights for designing pronunciation instruction that is both pedagogically sound and responsive to learners' needs.

| KEYWORDS

Systematic review (SR), pronunciation instruction, pronunciation pedagogy, pronunciation practice, pronunciation testing, practice with technology, EFL students, AI-generated speech, text-to-speech, pronunciation problems

| ARTICLE INFORMATION

ACCEPTED: 01 January 2026 **PUBLISHED:** 20 January 2026 **DOI:** 10.32996/jeltal.2026.8.1.2

1. Introduction

Pronunciation continues to be one of the most persistent challenges faced by second language (L2) learners. Research has documented a wide range of segmental and suprasegmental difficulties, as well as the cognitive, linguistic, and contextual factors that affect students' pronunciation performance. In response to the pronunciation challenges, numerous studies in the literature proposed a variety of instructional strategies, practice techniques, and technology-enhanced solutions aimed at improving students' pronunciation accuracy, fluency and intelligibility. Over the past decade, this growing interest has resulted in a substantial number of systematic reviews (SRs) and meta-analyses (MAs) that synthesize research on pronunciation pedagogy, technology-mediated pronunciation practice, and students' pronunciation acquisition.

The first group of SR and MA studies has focused on technology-enhanced pronunciation learning, particularly computer-assisted pronunciation training (CAPT), mobile-assisted language learning (MALL), automatic speech recognition (ASR), and artificial intelligence (AI) applications. These include SRs and MAs of AI applications for enhancing EFL learners' pronunciation (Aryanti & Santosa, 2024), technology-integrated blended teaching and its impact on English pronunciation accuracy among EFL learners in China (Ting, Kadir & Husain, 2025), the WordUp application (Wahyudi et al., 2025), CAPT (Amrate & Tsai, 2024; Jie, 2022; Mahdi & Al Khateeb, 2019; Almusharraf et al., 2024), ASR-based pronunciation training (Ngo et al., 2024), Arabic ASR (Dhouib et al., 2022), AI adoption in L2 learning (Almelhes, 2023), the effects of MALL on L2 pronunciation learning (Tseng et al., 2022), and exploring speech recognition and pronunciation accuracy detection techniques (Kaur & Bhathal, 2026). Other studies have examined mobile-assisted language learning and pronunciation practice (Metruk, 2024; Stoughton & Kang, 2024), application-based pronunciation practice (Suseno, 2023), shadowing for pronunciation (Whitworth, 2024; Whitworth & Rose, 2025), inclusive AI pronunciation tools (Meidyana et al., 2025), and pronunciation practice systems for speech-impaired children (Darejeh et al., 2019).

A second group of SRs and MAs addressed pronunciation instruction and instructional approaches, focusing on how pronunciation should be taught, teacher knowledge, instructional strategies, goals, and best practices. These include reviews of L2 pronunciation instruction effectiveness (Yağız et al., 2024), pedagogical content knowledge in Chinese as a foreign language (CFL) pronunciation teaching (Li et al., 2021), pronunciation instructional strategies in CFL (Li, 2018), foundations and future directions in pronunciation teaching (Kunova, 2025), non-native English-speaking teachers' (NNESTs) beliefs about pronunciation teaching (Wang et al., 2024), ideologies of nativeness and the intelligibility principle (Jeong & Lindemann, 2025), and goals and best practices in EFL pronunciation instruction (Ahn, 2013). Other reviews examined broader methodological and conceptual issues, such as leading scoping reviews on L2 pronunciation (Terrier et al., 2025) and corpus-supported perspectives on intelligibility (Jeong & Lindemann, 2025).

A third group of SRs and MAs synthesized measurable pronunciation outcomes, including intelligibility, comprehensibility, accentuation, pronunciation accuracy, and training effects. These studies have investigated relationships among fluency, intelligibility, comprehensibility, and accentuation (Chau & Huensch, 2025), the impact of high-variability phonetic training (Mahdi & Mohsen, 2024), explicit articulatory phonetics instruction (Vasquez & Aguirre, 2025), pronunciation challenges in specific national or curricular contexts such as Malaysia (Hamzah et al., 2015), Bangladesh (Qader et al., 2024), Iraq (Alrickaby, 2025). Further SR studies have explored phonetic systems and grapheme-to-phoneme correspondence, such as Hijaiyyah pronunciation among new learners (Mohamed et al., 2024) and grapheme-to-phoneme modelling for under-resourced languages (Irie et al., 2023), listening-based practice methods, including poetic style for pronunciation (Ariningsih, 2023) and podcasts for language learning (Ramirez, 2024).

Together, these SRs and MAs have significantly advanced our understanding of pronunciation instruction, technology-enhanced pronunciation learning, and student achievement. However, they share two important limitations. First, most reviews tend to focus on either pronunciation instruction or pronunciation practice, often within a single modality as Computer-Assisted Pronunciation Training (CAPT), Mobile-Assisted Language Learning (MALL), or Automatic Speech Recognition (ASR), rather than integrating traditional, digital, and AI-mediated approaches within a unified framework. Second, none of the existing SRs synthesize a coherent research program that covers segmental errors in specialized domains (e.g., biomedical terminology, proper nouns, silent consonants), prosodic and discourse-level phenomena in Arabic, AI-generated pronunciation errors, and technology-supported self-regulated practice using tools such as text-to- text-to-speech software, YouTube videos, mobile audiobooks, and digital multimedia labs, particularly in relation to Saudi EFL learners.

The author's research articles over two decades cover precisely these under-represented areas. They include pronunciation errors in AI-narrated Arabic YouTube videos; segmental pronunciation difficulties among Arab EFL learners and professionals; prosodic and discourse-level phenomena in spoken Arabic; and a wide range of technology-enhanced pronunciation and listening practices. The author's research program offers a unique opportunity to synthesize pronunciation instruction and pronunciation practice, whether traditional, digital, or AI-mediated, within a single SR. Therefore, the present study fills a significant gap in the literature by systematically reviewing the author's own research on pronunciation instruction and practice, across multiple contexts, modalities, and linguistic domains. The review aims to cover the thematic evolution of the author's research program, highlight its contributions to the broader pronunciation field, and identify pedagogical and technological implications, particularly for L2 learners and the broader L2 pronunciation field.

This study is significant because it addresses domains that no prior SR or MA has covered. Existing SRs focus on CAPT, MALL, ASR, High-Variability Phonetic Training (HVPT), AI tools, teacher beliefs, intelligibility, comprehensibility, accentedness, and pronunciation challenges in specific countries, but none examined AI-generated pronunciation errors. It analyzes pronunciation

errors in specialized domains as biomedical terminology, media terms, proper nouns, silent consonants, clipped borrowings, and casual speech segmentation. These are real-world pronunciation issues that SRs have completely ignored. It covers studies on prosody and discourse-level pronunciation in Arabic as intonational meanings of discourse markers, ambiguity in negative polar questions, and prosodic cues in spoken Arabic. No SR includes Arabic prosody or discourse-level pronunciation phenomena. It includes studies on technology for self-regulated pronunciation practice as text-to-speech for interpreting, YouTube for independent pronunciation practice, and mobile apps for listening, speaking and pronunciation practice. Prior SRs focus on instructional technology, not self-regulated practice environments. It covers studies on cognitive foundations of pronunciation. Early studies in this SR examined listening, decoding, and background knowledge. They connect pronunciation to comprehension and cognitive processing, a dimension that is absent from all SRs. The current SR integrates instruction + practice + AI + prosody + segmentals, whereas prior SRs are scattered: one SR on CAPT, one on MALL, another on teacher beliefs, a third on HVPT, another on vowel errors, another on Arabic ASR. On the contrary, the current SR is the first to integrate pronunciation instruction and practice, traditional methods, digital tools, AI-mediated materials, segmental errors, suprasegmental/prosodic aspects, discourse-level features, specialized domains, EFL + Arabic as a foreign language (AFL) + interpreting contexts. This breadth is unprecedented.

Moreover, this SR is the first to synthesize a 20-year research program. Most SRs summarize other people's work, whereas this SR synthesizes a coherent research trajectory, across multiple languages, multiple modalities, multiple learner populations, and multiple technological eras. This is academically rare and extremely valuable. It shows how this research evolved, how themes connect, how methods developed, how findings reinforce each other, how this work contributes to Saudi EFL and global L2 pronunciation research. This is the kind of SR that defines a scholar's legacy. It is directly relevant to Saudi EFL learners, Arab healthcare professionals, student-interpreters, Arabic phonology, Arabic prosody, Arabic discourse markers, AI-generated Arabic speech. This SR is the first to bring these together. It provides evidence-based insights, culturally relevant findings, pedagogically recommendations, technology-integrated solutions, and implications for curriculum design. This makes this SR uniquely valuable for Saudi universities, EFL programs, interpreting programs, healthcare English training, and AI-based educational tools, which Prior SRs rarely included. This research covers pre-digital era (2005), early digital tools (2018–2021), TTS and YouTube (2022), and AI-generated speech (2024–2025). It captures the historical evolution of pronunciation research which no other SR has done.

In short, the current study is significant because it is the first SR to integrate two decades of research on pronunciation instruction, pronunciation practice, technology-enhanced learning, and AI-generated speech across Arabic and English contexts, offering insights that no prior SR or MA has addressed.

Finally, the current SR is also part of a broader series of SR/MA projects by the author, including reviews of AI-assisted Arabic translation, linguistics, and pedagogy (Al-Jarf, 2026e); children's language acquisition in Saudi Arabia (Al-Jarf, 2026c); Arabic–English transliteration of personal names and public signage (Al-Jarf, 2026b); and a self-systematic review of English–Arabic and Arabic–English translation error studies (Al-Jarf, 2026a), Innovative Word Formation and Pluralization Processes in Arabic (Al-Jarf, 2026d).

2. Methodology

2.1 *Study Corpus*

The present SR is based on sixteen research articles published by the author between 2005 and 2025. Although the studies vary in focus, they share a unified methodological orientation: All examine pronunciation instruction an practice in L2 using corpus-based analysis and report quantitative data such as frequencies, percentages, and error rates. For the purpose of synthesis, the studies were grouped into five thematic clusters each representing a distinct strand of the author's research program on pronunciation instruction, pronunciation practice, and technology-mediated oral skills. Below is a description of each thematic cluster.

Cluster 1 — AI-Generated Speech and Pronunciation Errors

This cluster includes one study that examined pronunciation errors in Arabic YouTube videos narrated by AI (Al-Jarf, 2025g). The study highlights the emergence of AI-generated speech as a new pedagogical input source and identifies systematic segmental and suprasegmental inaccuracies that may affect Arabic-as-a-Foreign-Language (AFL) learners' listening comprehension and pronunciation acquisition.

Cluster 2 — Segmental Pronunciation Errors in Specialized Domains

This cluster includes six studies: Faulty consonant gemination in the pronunciation of English biomedical terms by Arab healthcare professionals (Al-Jarf, 2025c); splitting unsplittable foreign words in casual speech by EFL Arab Learners (Al-Jarf, 2025h); vowel pronunciation errors in English biomedical terminology by Arab healthcare professionals (Al-Jarf, 2025i);

pronunciation errors in English silent consonants by Saudi EFL Undergraduates (Al-Jarf, 2025e); proper noun pronunciation inaccuracies in English by educated Arabic speakers (Al-Jarf, 2022h); and student-interpreters' foreign proper noun pronunciation errors in English–Arabic and Arabic–English media discourse interpreting (Al-Jarf, 2022i). These studies focus on pronunciation challenges encountered by Arab EFL learners and professionals, including biomedical terminology, proper nouns, silent consonants, and phonotactic segmentation. The studies reveal persistent segmental difficulties that have direct implications for academic, professional, and communicative performance.

Cluster 3 — Prosody, Discourse and Spoken Arabic Phenomena

This cluster includes three studies addressing: intonational meanings of discourse markers in spoken colloquial Arabic (CA) (Al-Jarf, 2024d) ambiguity in Arabic negative polar questions (Al-Jarf, 2023b); and clipping of borrowings in Spoken Arabic (Al-Jarf, 2023c). Although centered on Arabic, these studies contribute to pronunciation pedagogy by illustrating how intonation, discourse markers, pragmatic ambiguity, and phonological processes in spoken Arabic shape meaning, interpretation, and communicative clarity.

Cluster 4 — Technology-Enhanced Pronunciation & Listening Practice

This cluster includes three studies as follows: Text-to-speech software for promoting EFL Freshman students' decoding skills and pronunciation accuracy (Al-Jarf, 2022k); YouTube videos as a resource for self-regulated pronunciation practice in EFL distance learning environments (Al-Jarf, 2022l); mobile technology and student autonomy in oral skill acquisition (Al-Jarf, 2012). Studies in this cluster investigate digital tools as text-to-speech software, YouTube, and mobile apps, as resources for pronunciation improvement, decoding skills, and self-regulated oral practice. These studies demonstrate the pedagogical value of technology-mediated environments in supporting autonomous pronunciation improvement.

Cluster 5 — Cognitive Foundations: Listening, Decoding & Pronunciation

This cluster includes three studies: effect of background knowledge on auditory comprehension in interpreting courses (Al-Jarf, 2018b); the effects of listening comprehension and decoding skills on spelling achievement of EFL freshman students (Al-Jarf, 2005); the relationship among spelling, listening and decoding skills in EFL freshman students (Al-Jarf, 2005). This cluster includes early foundational studies exploring the relationships among listening comprehension, decoding skills, background knowledge, and spelling. Although not exclusively pronunciation-focused, these studies provide essential cognitive factors for understanding how learners process and produce spoken language.

2.2 Eligibility (Inclusion & Exclusion) Criteria

To be included in the corpus, studies had to be authored by Reima Al-Jarf, published between 2005 and 2025, and contain extractable data relevant to Arabic–English pronunciation instruction or pronunciation practice. Because the dataset is a closed, author-bounded corpus, all publications were retrieved from the academic platforms listed in Section 2.4. No external database search was required. Based on these inclusion criteria, several groups of studies by the author were excluded because they fall outside the scope of Arabic–English pronunciation instruction and practice as follows:

- Duplicate studies as conference presentation that have similarly published articles were excluded such as pronunciation errors in AI-narrated Arabic YouTube videos (Al-Jarf, 2025f); can students learning Arabic as a foreign language use Arabic YouTube videos narrated by Artificial Intelligence (AI) for listening practice (Al-Jarf, 2025b) and effect of background knowledge on auditory comprehension in interpreting courses (Al- Jarf, 2018a).

- Transliteration studies that focus on Arabic–English transliteration conventions, orthographic variation, or social-media spelling practices rather than pronunciation instruction. Excluded studies include: Arabic transliteration of borrowed English nouns with /g/ by Artificial Intelligence (AI) (Al-Jarf, 2025a); semantic and syntactic anomalies of Arabic-transliterated compound shop names in Saudi Arabia (Al-Jarf, 2023g); absence of vowels in the English spelling of Arabic personal names on social media (Al-Jarf, 2023a); English spelling of the glottal stop and voiced pharyngeal fricative in Arabic personal names by educated Arabs (Al-Jarf, 2023f). English spelling of Arabic compound personal names by educated Arabs (Al-Jarf, 2023d); English transliteration of Arabic personal names with the definite article /al/ (Al-Jarf, 2022e); deviant Arabic transliterations of foreign shop names in Saudi Arabia and decoding problems among shoppers (Al-Jarf, 2022b); gemination errors in Arabic-English transliteration of personal names (Al-Jarf, 2022f); and variant transliterations of the same Arabic personal names (Al-Jarf, 2022m).

- Spelling and phoneme-grapheme correspondence studies that examine spelling strategies, orthographic errors, or decoding skills rather than pronunciation instruction were excluded: EFL freshman students' difficulties with phoneme-grapheme relationships (Al-Jarf, 2019); teaching spelling with mind-mapping software (Al-Jarf, 2011); spelling error corpora in EFL (Al-Jarf, 2010); auditory and visual problems of good and poor EFL college spellers (Al-Jarf, 2009); listening-spelling

strategies in EFL Arab college students (Al-Jarf, 2008b); phonological and orthographic problems in EFL college spelling (Al-Jarf, 2008c); sources of spelling errors in EFL Arab college students (Al-Jarf, 2008d); faulty strategies of EFL freshman spellers, Saudi Arabia (Al-Jarf, 2007a); the effects of listening comprehension and decoding skills on spelling achievement of EFL freshman students (Al-Jarf, 2005); the relationship among spelling, listening and decoding skills in EFL freshman students (Al-Jarf, 2005); listening-spelling strategies of freshmen students (Al-Jarf, 1999).

- Studies on listening and speaking that address general oral-skill development, speaking assessment, or reading–speaking integration rather than pronunciation instruction per se were excluded: developing students' global awareness in EFL reading and speaking (Al-Jarf, 2022c); EFL college students' and instructors' preferred method of speaking assessment in the Saudi context (Al-Jarf, 2021a); EFL speaking practice in distance learning during the coronavirus pandemic 2020–2021 (Al-Jarf, 2021b); integrating participation goals in writing activities for EFL college students (Al-Jarf, 2021d); enhancing reading and speaking skills in EFL through multicultural children's short stories (Al-Jarf, 2015a); integrating global themes in EFL speaking instruction (Al-Jarf, 2008d); the effect of TBLT on students' speaking ability (Al-Jarf, 2007b); TED talks as a listening resource in EFL college classrooms (Al-Jarf, R. (2021g); integrating TED lectures in EFL college listening practice (Al-Jarf, 2020a); and mobile audiobooks, listening comprehension and EFL college students (Al-Jarf, 2021e).

- Studies that focus on interpreting competence, interpreting pedagogy, or directionality rather than pronunciation instruction were excluded: comparative study of directionality in English-Arabic and Arabic-English interpreting competence of student interpreters (Al-Jarf, 2022d); integrating current global events and technology in interpreting practice (Al-Jarf, 2022g); teaching interpreting for tourism purposes (Al-Jarf, 2021f); feasibility of digital multimedia language labs for interpreting instruction (Al-Jarf, 2021c); text-to-speech software as a resource for independent interpreting practice (Al Jarf, 2022j); what instructors and students should know about interpreting problems (Al-Jarf, 2015b); how to teach liaison interpreting to beginners (Al-Jarf, 2007c); and bridging the gap between teacher and learner in liaison interpreting (Al-Jarf, 2000).

- Studies that deal with pronunciation partially as in mobile apps in the EFL college classroom (Al Jarf, 2020b); grammar podcasts for ESL college students in distance learning (Al Jarf, 2023e); a multiple-associations approach to teaching technical terms in English for specific purposes courses (Al-Jarf, (2022a); testing multiple vocabulary associations for effective long-term learning (Al-Jarf, 2023h); and making connections in vocabulary instruction (Al-Jarf, 2006).

2.3 Corpus Characteristics

The final corpus consisted of sixteen studies authored by Reima Al-Jarf between 2005 and 2025. Because the dataset represents a closed research program spanning two decades, the corpus is both comprehensive and internally coherent, reflecting the author's sustained scholarly trajectory in pronunciation, phonology, technology-enhanced oral skills, and AI-mediated speech. The included studies vary in methodological design, encompassing qualitative analyses, quantitative error counts and percentages, technology-based intervention studies, and descriptive investigations.

All included articles share a common focus on pronunciation accuracy, phonological processing, listening–decoding relationships, or technology-supported pronunciation practice. Across the corpus, the studies collectively examine segmental pronunciation errors (e.g., vowels, consonants, gemination, silent letters, proper nouns); suprasegmental and discourse-level features (e.g., intonation, discourse markers, pragmatic ambiguity); technology-mediated pronunciation practice (e.g., text-to-speech software, YouTube, MP3 lessons); AI-generated speech as an emerging pedagogical input source; and cognitive foundations of pronunciation (listening comprehension and decoding). The corpus covers a wide range of learner populations (EFL undergraduates, healthcare professionals, interpreting students, distance-learning learners, and general adult EFL learners). The studies also cover multiple linguistic contexts (biomedical terminology, proper nouns, CA discourse, and AI-generated narration).

To facilitate synthesis, the sixteen studies were organized into five thematic clusters, each representing a distinct dimension of the author's research program (See 2.1 above). Together, these clusters provide a comprehensive overview of the author's contributions to pronunciation pedagogy, phonological analysis, and technology-supported oral-skills development. The corpus reflects a longitudinal, multi-modal exploration of pronunciation challenges and instructional solutions relevant to Arabic–English language learning.

2.4 Information Sources

The information sources were limited to platforms that index the author's complete scholarly output. No external database search was required, as the aim was not to identify all studies on a broad topic, but to synthesize all pronunciation-related studies within a single, self-contained research program. All records were retrieved from publicly accessible academic databases

in which the author's publications are fully archived. These include Google Scholar, ResearchGate, Semantic Scholar, Academia.edu, SSRN, ERIC, EBSCO, ProQuest, and one article indexed in Scopus, Web of Science (e.g., *mobile technology and student autonomy in oral skill acquisition*, Al-Jarf, 2012), institutional repositories such as King Saud University repository, and publisher platforms. Together, these sources provide complete coverage of the author's publications across journals, conferences, book chapters, and digital repositories. All included and excluded studies were verified manually to ensure accuracy, remove duplicates, and confirm alignment with the eligibility criteria described in Section 2.2.

2.5 Data Extraction

For every study, the following data were extracted from the full text of the article: publication year; research focus (e.g., segmental errors, suprasegmental features, AI-generated speech, technology-enhanced pronunciation practice); participant characteristics (e.g., EFL undergraduates, healthcare professionals, student-interpreters); linguistic targets (e.g., vowels, consonants, proper nouns, discourse markers); methodological approach (qualitative analysis, corpus-based mapping, descriptive error analysis, technology-mediated intervention); data sources (e.g., learner recordings, spontaneous speech, AI-generated narration, YouTube materials, TTS output, classroom tasks); and key findings relevant to pronunciation accuracy, error patterns, or pedagogical implications. These elements are necessary for thematic synthesis and cluster-level comparison rather than effect-size calculation, as the corpus consists of qualitative analyses, descriptive studies, and technology-enhanced pronunciation investigations rather than controlled experimental trials.

Coding was made manually to preserve conceptual accuracy and to ensure that each study was classified according to its primary pronunciation-related contribution. Any study that contained overlapping themes (e.g., pronunciation + technology, or pronunciation + discourse) was assigned to the cluster that best reflected its central research question. All extracted data were entered into a structured matrix to ensure consistency across studies and to facilitate comparison within and across the five thematic clusters. Because the corpus is limited to a single author's research program, terminology, methodological framing, and analytical categories were highly consistent, reducing coding discrepancies.

The above systematic extraction process ensured that all included studies were analyzed using uniform criteria, enabling a coherent synthesis of pronunciation-related findings across two decades of research.

2.6 Data Synthesis

Data synthesis combined qualitative, narrative, and cluster-based approaches, reflecting the descriptive and heterogeneous nature of the included studies. Because the corpus consists of pronunciation-focused investigations authored by a single researcher over a twenty-year period, the synthesis emphasized conceptual integration, pattern identification, and cross-study comparison. The synthesis proceeded in three stages: (i) All studies were first grouped into five thematic clusters based on their focus, linguistic targets, and methodological orientation (See section 2.1 above). This classification enabled the review to synthesize findings within conceptually unified domains while preserving the distinct contributions of each study. (ii) Within each cluster, studies were compared according to linguistic features examined (segmental, suprasegmental, discourse-level), learner populations and contexts, data sources (AI narration, learner speech, YouTube videos, TTS output, classroom tasks), methodological procedures (error analysis, corpus mapping, qualitative interpretation, technology-based intervention), recurring error patterns and pedagogical implications. (iii) Findings were synthesized across clusters to highlight broader patterns in the author's pronunciation research, to focus on convergence of error types across different linguistic domains, the interaction between cognitive processes (listening, decoding) and pronunciation accuracy, cross-context consistency in learner difficulties, pedagogical recommendations supported by multiple studies and the role of technology and AI as emerging pronunciation input sources. This cross-cluster synthesis provided a coherent picture of pronunciation challenges and instructional solutions relevant to Arabic–English learners because the corpus is methodologically aligned and conceptually interconnected.

2.7 PRISMA Flow Description

Because this review is based on a closed, predefined corpus consisting exclusively of sixteen studies published by the author between 2005 and 2025, the PRISMA flow reflects a simplified identification and screening process. All publications within this time frame were retrieved from the academic platforms listed in Section 2.4 and manually screened for relevance. After removing duplicates, all records were assessed against the eligibility criteria. Studies were excluded if they focused on transliteration, spelling and phoneme–grapheme correspondence, general listening or speaking skills, or interpreting, as these topics fall outside the scope of Arabic–English pronunciation instruction and practice. Following full-text evaluation, only studies directly addressing pronunciation-related phenomena, such as segmental and suprasegmental features, AI-generated speech, technology-enhanced pronunciation practice, and cognitive foundations of pronunciation, were retained. These studies were subsequently organized into five thematic clusters for synthesis. The PRISMA flow therefore documents the progression from identification of all publications within the author-bounded corpus, through screening and eligibility assessment, to the final inclusion of studies contributing directly to pronunciation instruction and practice.

3. Results

3.1 Overview

The sixteen studies in this SR examine segmental and suprasegmental features, AI-generated speech, technology-enhanced pronunciation practice, and the cognitive processes underlying listening and decoding. The sixteen studies included in this RS, were organized into 5 thematic clusters representing the major strands of the author's research program: (1) AI-generated speech and pronunciation errors; (2) segmental pronunciation errors in specialized domains; (3) prosody, discourse, and spoken-Arabic phenomena; (4) technology- enhanced pronunciation and listening practice; and (5) cognitive foundations of pronunciation, listening, and decoding. The following sections present the findings within each cluster and highlight the key contributions of the author's work to the field of pronunciation instruction and practice.

3.2 Study Characteristics

Cluster 1 — AI-Generated Speech and Pronunciation Errors:

Study 1: Pronunciation errors in Arabic YouTube videos narrated by AI (Al-Jarf, 2025g)

This study analyzes pronunciation errors in Arabic YouTube videos narrated by AI, focusing on how the absence of short-vowel diacritics affects the accuracy of AI-generated speech. Although AI narrators produce natural-sounding voices with appropriate intonation and no grammatical or syntactic errors, they consistently mispronounce words whose meanings and pronunciations depend on short vowels that are not represented in standard Arabic orthography. The analysis revealed recurrent errors in short-vowel diacritics, particularly /a/, /u/, /i/; homographs, where identical spellings correspond to different meanings and pronunciations; suffixal inflections, especially تاء التأنيث and person/tense markers (e.g., كُتِبَتْ، كتبتِ، كتبتٌ، كتبتَ); and context-dependent pronunciation, where AI failed to match the correct vocalization to the surrounding syntactic or semantic cues. These errors result in semantic ambiguity and reduced intelligibility for learners of Arabic as a foreign language (AFL), and noticeable distortion for native speakers. The study also explains why AI systems can maintain native intonation mispronouncing individual words, with prosody generated algorithmically, whereas vowel assignment requiriring lexical and contextual interpretation that current models do not reliably perform.

Cluster 2 — Segmental Pronunciation Errors in Specialized Domains

Study 2: Faulty consonant gemination in the pronunciation of English biomedical terms by Arab healthcare professionals (Al-Jarf, 2025c)

Healthcare professionals tend to geminate consonants such as /l, r, n, b, t, p/ in words like *penicillin, cannula, collagen, millimeter,* and *heart attack*. These errors stem from L1 phonological transfer, misinterpretation of English orthography, and Arabic phonotactic habits, where gemination is phonemic and contrastive. Participants also geminated single final consonants in words such as *up, cut, shut,* and *gel*, reflecting repair strategies influenced by Arabic monosyllabic words ending in true geminates. By contrast, they did not geminate double letters in terms like *Accu-Check, mammography,* or *capillaries*, where the doubled consonants occur in unstressed or syllable-boundary positions that do not trigger gemination in Arabic. Overall, the study demonstrates predictable L1-driven patterns in the pronunciation of English biomedical terms and highlights how Arabic phonotactics shape consonant lengthening in professional healthcare communication.

Study 3: Vowel Pronunciation Errors in English Biomedical Terminology by Arab Healthcare Professionals (Al-Jarf, 2025i)

Arab healthcare professionals such as doctors, dentists, pharmacists, physical therapists, nutritionists, lab technicians, radiologists mispronounce vowels in biomedical terms such as pronouncing terms with y & i as /ai/ in (*acetyl, methyl, vertigo, cervical*) as in the letter name in English, relate them to "*analyze*", overgeneralize the pronunciation of i in *combine* to *combination*, pronounce the suffix -gia in neuralgia & metatarsalgia orthographically as /dʒjə/ instead of /dʒə/; lengthen "-in", "-on" & "-ol" ending to /iːn/, /ɔːn/ & /ɔːl/ in *Aspirin, Insulin, Relaxon, Parafon* due to Arabic phonetic tendencies, overcorrection and by analogy with *caffeine*. /eu/ in *Euthyrox* and *Eucarbon* is pronounced /ʊ/ not /yʊ/. By contrast the eu in *Neuroton* is pronounced /ju/ or /jʊ/ instead of /ʊ/ or /u/. In *diet*, the triphthong /ˈdaɪ.ət/ or /ˈdaɪ.ɪt/ is reduced to the diphthong /ai/ (/daɪt/). The vowels a, e, o in *Galvus & Omega* are shortened. They changed the vowel in "*Rapidus,*" "*Centrum,*" "*Maximum*," from /ə/ sound (schwa) to /ʊ/. Vowel mispronunciations can be attributed to orthographic influence (*cervical*), overgeneralization from related words (*combine*), overgeneralization of familiar patterns to new biomedical terms as the analogy between *amylase* and *analyze* and applying phonological interference, leading to vowel shifts.

Study 4: Splitting Unsplittable Foreign Words in Casual Speech by EFL Arab Learners (Al-Jarf, 2025h)

Arabic-speaking EFL learners split unsplittable foreign long words in the flow of speech into two sub-parts as in *Skype* > Sky + pe, *Kaspersky* > Kasper + sky, *Swarovski* > Swaro+viski, *Google* > Go + gil, *vegetable* > vege + table, *marshmallow* > Marsh + mello, *Michigan* > Mit + shigan, *Wednesday* > Wednes + day, *manipulated* > manu + plated and so on. In segmenting long words, Arab learners often rely on the words' written form, treating unfamiliar long words as consisting of familiar parts and pronouncing them as if they were two words, with a slight pause between both parts. They rearrange consonant clusters based on their Arabic (L1) phonotactic constraints, insert a vowel to break the clusters, and stress the penultimate syllable in the second

part. Some faulty word segmentation in the sample is based on cross-linguistic lexical associations, where segments evoke meaningful words in their native language (Arabic) as Swaro سوار meaning (*bracelet*).

Study 5: Mapping Pronunciation Errors in English Silent Consonants by Saudi EFL Undergraduates (Al-Jarf, 2025e);
Saudi EFL students had varying degrees of difficulties with silent d (27%) (*handsome, grandmother, windmill*), Silent t (21%) (*castle, soften, wrestling*); Silent b (14%) (*bomb, bombing, plumber*); Silent l (8%) (*salmon, almond*); Silent h (9.5%) (*honor, honest*); Silent s (7%) (*island, Illinois*); and Silent n (6%) (*autumn, mnemonic*). They had less difficulty with Silent k (3%) (*knee, knife*), Silent gh (*light, through*) and Silent g (*sign, design, foreign*). Pronouncing silent consonants in English words can be attributed to transfer from Arabic which is a phonetic language that lacks true silent consonants. Many students rely on spelling to determine pronunciation. Also, the students lack exposure to the pronunciation of English native speakers.

Study 6: Proper Noun Pronunciation Inaccuracies in English by Educated Arabic Speakers (Al-Jarf, 2022h)
In pronouncing Proper Nouns used in English, Arabic speakers have the following problems:(i) mispronouncing English vowels in *Google, Moodle, Uber, Nixon, London*; (ii) replacing consonants absent in L1 (pv) by their equivalents (*bebsi, jafa*); (iii) geminating consonants in city and country names (*Peking; Venezuela, Minnesota*); (iv); inserting a vowel in consonant clusters in Proper Nouns and acronyms (*Zelinsky, Lugansk, SNAS, GMC*,); (v) breaking words into two sub-words (Kasper+ sky, Sky+ pe); (vi) pronouncing words the way they are spelled (*Nazi, Nike, Huawei*, *Hyundai, Wednesday*); and (vii) transferring Arabic stress rules to English words (*McDonald, Mayflower*). These pronunciation errors are attributed to transfer from Standard Arabic (SA) or the students' local dialect, insufficient mastery of English pronunciation rules, phonics and phone-grapheme correspondences and lack of knowledge of the differences between English and Arabic phonology.

Study 7: Student-interpreters' foreign proper noun pronunciation errors in English–Arabic and Arabic–English Media discourse interpreting (Al-Jarf, 2022i)
Student interpreters have difficulty identifying and discriminating one or more phonemes in unfamiliar foreign Proper Nouns. They make up nonsense words that rhyme with the unfamiliar source words (**Dagos, *Dados, *Dabos* for *Davos; *Scinavia* for *Scandinavia*; and **NADO for NATO)*. They deleted part of the Proper Noun (**Buja* instead of *Abuja*, **Izheimer, *Bloombe*r). Phonemes were changed and substituted by a longer or shorter vowel, by another consonant or another syllable in *Dracula /dracola/, *foks fagon, Ukraine /ʊkrɜːrɪə/.* The Arabic pronunciation was retained and overgeneralized in *Eiffel Tower* /iːfəl/ or /iːvəl/, **Ardoghan*. A vowel was inserted to break the consonant clusters in **Beligrade, *Barazil, *Shangahai*. Syllables were reversed in **Serbrenica* and **ALESCO*. Most pronunciation errors are attributed to lack of knowledge of foreign Proper Nouns commonly occurring in the media.

Cluster 3 — Prosody, Discourse & Spoken Arabic Phenomena
Study 8: Intonational meanings of discourse markers in spoken colloquial Arabic (Al-Jarf, 2024d)
This study investigated the types of meanings and pragmatic functions that the discourse markers (طيب /Tayyib/ O.K, خلاص /xala:S/ (that's it), إن شاء الله InshaAllah (God willing), ما قصّرت /ma gaSSart/ & ما قصّرتي /ma gaSSarti/ (much appreciated), لا لا /la: la:/ (no…no), يا ستي /ya sitti/ (ma'am), يا سلام /ya sala:m/ (wow), يا عيني /ya ʔeyni/, يا مسهل /ya: msahhil/ (asking God for making things easy), and يا ساتر /ya sa:ter/ (Oh My God) have when each is uttered with different intonation patterns in spoken CA. Twenty student-translators received training in uttering discourse markers with different intonations, identifying the meaning and/or purpose conveyed by each intonation, then they performed an elicitation and a judgment/interpretation task in which they were required to pronounce each discourse marker out loud with as variety of intonations and identify the meaning conveyed by each. Results showed that each discourse marker has a variety of meanings and pragmatic functions when uttered with different intonations. The context makes it clear which meaning each intonation implies.

Study 9: Ambiguity in Arabic negative polar questions (Al-Jarf, 2023b)
In this study, a sample of negative polar questions in Hijazi Arabic (HA) was uttered with different rising intonations by a sample of students enrolled in a Semantics and Pragmatics course. The student informants were asked about the meaning conveyed by each intonation of the same negative polar question. Results showed that ما شريت فستان الأسبوع الماضي؟ Didn't you buy a dress last week?" is a negative polar question formed with a change in intonations. It is ambiguous and may render the following meanings: (i) a neutral question about whether she bought the dress or not, replying with the truth-value of the situation, or is replying to the polarity used in the question. The answer would be either "Yes" or "no", or an echo answer: "Yes I bought it" or "No I didn't buy it"; (ii) a confirmation question to which the reply is "yes" only; (iii) a confirmation question to which the reply is "no" only; (iv) disapproval: "Didn't you buy a dress last week? Why do you want to by another dress? (v) an exclamation: Wow! You have bought a new dress, although you bought one last week! What a surprise! The context makes it clear which meaning each intonation implies.

Study 10: Clipping of borrowings in spoken Arabic (Al-Jarf, 2023c)

Results revealed five types of borrowings: (1) back clipping of single words and compound (70%) (*centimeter* سانتي, *hypermarket* هايبر, *Facebook* فيس, *WhatsApp* واتس*);* (2) fore-clipping (10%) as in *album* البوم*;* (3) medial clipping (11%) (*Alzheimer* ازهايمر, *cinema* السيما, *radio*)رادو*;* (4) complex clipping with phonological changes (9%) (rickshaw >ricksha, Hollywood >hilyood); and (v) crasis (*workshop* ورشة عمل, *screwdriver* سكروب). Such clippings are spontaneously created and used by educated and uneducated Arabs and those who do not know English. Students majoring in English, medicine, and pharmacy create their own clippings: *Style (Stylistics); Semantic (Semantics); Contrastive (Contrastive Analysis); mid (midterm), cause (because), lap (laptop), pharma (pharmacy), ophtha (ophthalmology).* The students asserted that clipped forms are *"cute, easy, and everybody is doing it"*. They use them with other students in the same major. Instructors should draw students' attention to student-created clippings that are ungrammatical, and to Arabic clipped borrowings that can only be used in spoken Arabic but not in English formal writing. Socio- and psycholinguistic reasons for clipped borrowings are given.

Cluster 4 — Technology-enhanced pronunciation and listening practice

Study 11: Text-to-speech software for promoting EFL students' decoding skills and pronunciation accuracy (Al-Jarf, 2022)

A Quasi-experimental study with two freshman groups (experimental vs. control) enrolled in Vocabulary I and Reading I courses was conducted. Both groups received identical in-class instruction. The experimental group used NaturalReader text-to-speech software (TTS) for weekly decoding and oral reading practice over 12 weeks. A Recognition (vocabulary) test, Production (oral reading) test, Weekly intermediate tests, usage frequency (number of lessons/texts practiced + practice time), post-treatment attitude questionnaire were used to measure students' progress. Results showed that the experimental group scored significantly higher than the control group on the decoding (recognition) and pronunciation (oral reading) posttests. Improvement in decoding and pronunciation was slow but steady, with significant gains noted after 8 and 12 weeks of the TTS practice. Vocabulary knowledge did not improve significantly. Strong positive correlations were found between recognition and production posttest scores (r = .89), and usage frequency and posttest performance (r = .84–.85). Students who practiced more lessons and spent more time with NaturalReader achieved higher decoding and pronunciation proficiency. Attitudes toward TTS practice were overwhelmingly positive. The students reported increased fluency, confidence, and enjoyment. The study provides empirical evidence on how TTS supports decoding, pronunciation accuracy, and oral reading fluency, with clear statistical outcomes and longitudinal improvement patterns.

Study 12: YouTube videos as a resource for self-regulated pronunciation practice in EFL distance learning environments (Al-Jarf, 2022)

Many EFL students have limited opportunities to listen to native speakers, practice English out of class, have problems in listening comprehension, oral expression and lack oral fluency. This article proposes using YouTube videos to improve EFL students' pronunciation. It aims to show the following: (i) How YouTube videos can be integrated in EFL instruction to teach pronunciation; (ii) the advantages of integrating YouTube videos: They are free, provide variety of topics, speakers, difficulty level, can be easily downloaded to the laptop or mobile phone and can be viewed anywhere and anytime; (iii) show how examples YouTube videos that target a specific pronunciation skill can be located; (iv) criteria for selecting online videos such as the video length in minutes, topic familiarity, difficulty level, speed of the speakers, students' proficiency level, and students' interests; (v) pronunciation subskills that can be developed through supplementary YouTube videos; (vi) teaching and learning with YouTube videos before watching a video, while watching the video and after watching the video.

Study 13: Mobile technology and student autonomy in oral skill acquisition (Al-Jarf, 2012)

Results showed that students who used the MP3 self-study program made significantly higher gains in oral skill acquisition than those who relied on the textbook only. After eight weeks, the experimental group's oral fluency scores improved by 32%, compared to 12% in the control group. Listening comprehension also increased substantially: the experimental group's mean listening score rose from 41% to 68%, while the control group improved from 40% to 52%. Auditory discrimination accuracy increased by 29% in the experimental group versus 10% in the control group. Self-reported autonomy scores increased from 2.4 to 4.1 in the experimental group, compared to a modest rise from 2.5 to 2.9 in the control group. Usage logs showed that students in the experimental group completed 90 lessons and listened to 900 audio files, averaging 110 minutes of independent practice per week. 78% of the students reported that MP3 lessons made them more willing to practice English outside class, and 72% indicated that the MP3 lessons improved their confidence in speaking. Overall, Results demonstrated that mobile-assisted self-study significantly enhanced oral fluency, listening comprehension, auditory discrimination, and learner autonomy, outperforming traditional textbook-only instruction across all measured variables.

Cluster 5 — Cognitive Foundations: Listening, Decoding & Pronunciation

Study 14: Effect of background knowledge on auditory comprehension in interpreting courses (Al-Jarf, 2018b)

The results showed a strong correlation between interpreting accuracy and vocabulary knowledge (r = .58, p < .01). Students had difficulty discriminating phonemes in unfamiliar Proper Nouns and technical terms, producing nonsense substitutes (75%),

rhyming approximations (14%), and reductions (10%) such as *Dignero* for *Rio de Janeiro*, *Al go* for *Al Gore*, and *Buja* for *Abuja*. Semantic errors were also frequent: *WHO* was interpreted as a question word by 66% of students, *FAO* was misrendered by 40%, and polysemous items such as *inflated* were assigned meanings inappropriate to context. Error sources included inadequate L1 background knowledge, limited EFL vocabulary, and difficulty selecting the correct meaning of polysemous words. Students relied on compensatory strategies such as nonsense-word creation (50%), sound analogy, literal translation, overgeneralization of Arabic forms (Athina), and use of erroneous equivalents (e.g., *EU → Soviet Union*).

Study 15 & 16: The effects of listening comprehension and decoding skills on spelling achievement of EFL students (Al-Jarf, 2005b) & The relationship among spelling, listening and decoding skills in EFL students (Al-Jarf, 2005a)

EFL freshman students showed poor performance across spelling, listening and decoding skills. On average, they misspelled 41.5% of the dictated words, answered 49.5% of the listening items correctly, and decoded 52% of the items on the test accurately, with a wide score variability in all measures. ANOVA results revealed a significant difference among the three mean scores ($F = 4.73$, $p < .05$), with listening (52.8%), decoding (51.6%), and spelling (39.8%). The was a strong a strong significant negative correlation between misspellings and listening comprehension ($r = –0.75$) and between misspellings and decoding ($r = –0.73$), indicating that students with better listening and decoding skills made fewer spelling errors. A positive correlation was also found between listening and decoding ($r = 0.65$), showing that the two skills tend to develop together. Overall, the results demonstrate that poor spelling performance is closely linked to poor listening comprehension and decoding ability. This means that good spelling ability in EFL is related to good listening comprehension and good decoding skills. The better the listening comprehension and decoding abilities, the fewer the spelling errors. When listening comprehension and decoding skills are poor, spelling ability is also poor.

4. Discussion

4.1 Meta-Conclusion

Across the five thematic clusters, the findings of this SR reveal a coherent research program that consistently demonstrates how pronunciation accuracy in Arabic–English contexts is shaped by the interaction of phonological knowledge, decoding ability, background knowledge, and the quality of input, whether human or AI-generated. Studies on AI speech show that synthetic audio introduces systematic segmental and suprasegmental errors that can mislead learners. Research on Arab EFL learners and professionals highlights the effect of L1 such as consonant gemination, faulty segmentation of long words, and mispronunciation of foreign Proper Nouns. Work on interpreting and listening comprehension further shows that limited lexical knowledge, inadequate background knowledge, and poor decoding skills significantly impair auditory discrimination and meaning transfer. Collectively, these studies converge on a central conclusion: pronunciation performance is not an isolated skill but a composite outcome of phonological processing, lexical access, decoding efficiency, and contextual knowledge, all of which interact dynamically in real-time speech perception and production. The author's two-decade research therefore provides an integrated account of how learners process, interpret, and produce spoken language, and how both human and AI-generated input can facilitate or hinder accurate pronunciation.

4.2 *Meta-Interpretation.*

Taken together, the findings across the five clusters reveal a unified pattern: pronunciation accuracy in Arabic–English contexts is shaped less by isolated phonetic skills and more by the interaction of phonological processing, lexical access, decoding efficiency, and background knowledge. The studies consistently show that when any one of these components is weak, whether it is the ability to discriminate phonemes, retrieve lexical items, decode unfamiliar forms, or draw on contextual knowledge, pronunciation accuracy deteriorates in predictable ways. AI-generated speech introduces its own layer of complexity: although prosodically fluent, synthetic voices often mispronounce context-dependent forms such as homographs and diacritic-sensitive Arabic words, creating misleading input for learners who rely on AI as a pronunciation model. Similarly, learners' segmental errors, such as gemination, faulty word segmentation, or mispronunciation of Proper Nouns, reflect deep L1 phonotactic influences and limited exposure to accurate models. Interpreting studies further demonstrate that pronunciation errors are rarely random; they emerge from cognitive overload, gaps in world knowledge, and limited access to the correct meaning of polysemous or specialized terms. Across all clusters, the evidence points to a shared mechanism: when learners cannot map sound to meaning efficiently, they compensate through approximation strategies, including nonsense forms, analogies, literal translation, or overgeneralization. These compensatory behaviors highlight the cognitive pressures involved in real-time speech processing and underscore the need for pronunciation instruction that integrates phonological training with decoding, vocabulary development, and contextual knowledge.

4.3 *Cross-Cutting Insights*

Across the five thematic clusters, several cross-cutting insights emerge that illuminate the shared mechanisms underlying pronunciation, listening, and decoding performance in Arabic–English contexts. First, the studies consistently show that L1 phonological transfer is a strong influence across learner groups and task types. Whether in gemination, faulty segmentation of

long words, or mispronunciation of foreign Proper Nouns, learners rely on Arabic phonotactic patterns when processing unfamiliar English forms. Second, lexical access and background knowledge repeatedly seem to be decisive factors: when learners are unfamiliar with names, technical terms, or institutional designations, they resort to approximation strategies that distort both sound and meaning. This pattern appears in interpreting, in spelling–listening–decoding relationships, and in the processing of AI-generated speech. Third, the findings highlight the central role of decoding efficiency, the ability to map orthographic, phonological, and semantic cues onto one another. Weak decoding skills lead to errors in spelling, misinterpretation of media discourse, and difficulty assigning correct vowel patterns in Arabic or stress patterns in English. Fourth, the studies collectively demonstrate that input quality matters. AI-generated speech, despite its fluent intonation, introduces systematic pronunciation errors that can mislead learners, while inconsistent exposure to accurate models contributes to persistent segmental and suprasegmental inaccuracies. Finally, across all clusters, learners employ similar compensatory strategies, nonsense forms, rhyming analogies, literal translation, overgeneralization, and vowel insertion, revealing shared cognitive pressures in processing. Together, these insights show that pronunciation accuracy is shaped by an interconnected system of phonological, lexical, cognitive, and contextual factors rather than isolated skill components.

4.4 *Causes of Pronunciation Problems*

Across the studies included in this review, several interrelated causes of pronunciation problems emerged. The first factor is L1 phonological transfer, where learners rely on Arabic phonotactic patterns in processing English or foreign names. This transfer appears in gemination, vowel insertion, faulty segmentation of long words, and mispronunciation of unfamiliar Proper Nouns. A second major cause is limited lexical knowledge, particularly of technical terms, international organizations, political designations, and foreign place names. When learners cannot access the correct lexical item, they approximate through rhyming forms, nonsense substitutes, or literal translation, which distorts both sound and meaning. A third cause is poor decoding skills, reflected in difficulty identifying orthographic, phonological, and semantic cues onto one another. Poor decoding contributes to spelling errors, misinterpretation of media discourse, and inaccurate assignment of stress, vowel quality, or consonant length. A fourth cause is insufficient background knowledge, especially in interpreting tasks where comprehension depends on familiarity with global events, institutions, and specialized terminology. Without this knowledge, learners cannot anchor unfamiliar sounds to meaningful referents. Finally, the studies highlight the role of input quality, particularly the influence of AI-generated speech. Although prosodically fluent, AI narration often mispronounces homographs, diacritic-dependent Arabic words, and context-sensitive forms, providing learners with inaccurate models. Together, these causes show that pronunciation problems arise, not from isolated phonetic weaknesses, but from the combined effects of phonological transfer, lexical gaps, decoding limitations, contextual unfamiliarity, and inconsistent exposure to accurate input.

4.5 *Implications*

The cross-cluster findings of this review carry several important implications for pronunciation research and pedagogy in Arabic–English contexts. First, the consistent influence of L1 phonological transfer suggests that pronunciation instruction cannot rely solely on corrective feedback or isolated drills; it must explicitly address how Arabic phonotactics shape learners' perception and production of English sounds. This includes targeted work on consonant clusters, stress patterns, vowel quality, and the avoidance of over-gemination. Second, the strong links between lexical knowledge, decoding ability, and pronunciation accuracy indicate that pronunciation should be integrated with vocabulary development and decoding training rather than treated as a standalone skill. Learners who cannot access the correct lexical item or decode unfamiliar forms are more likely to mispronounce, approximate, or distort both sound and meaning.

4.6 *Positioning This Work Within the Global Pronunciation Research*

The body of research synthesized in this review aligns with, and meaningfully extends, global scholarship on second-language pronunciation, speech perception, and cross-linguistic phonological transfer. Internationally, pronunciation research has increasingly emphasized the role of L1 phonotactics, lexical access, and decoding skills in shaping learners' ability to perceive and produce L2 sounds. The findings across the five clusters reinforce these global trends while offering a uniquely detailed account of how these mechanisms operate in Arabic–English contexts, a linguistic pairing that remains underrepresented in mainstream pronunciation literature.

This study also contributes to global research by emphasizing the role of background knowledge, as familiarity with Proper Nouns, institutional terminology, and global events, in real-time speech processing. While international studies often focus on phonetic accuracy or intelligibility, the present corpus demonstrates that pronunciation errors frequently stem from knowledge-based constraints, not only phonetic ones. This positions the author's research at the intersection of pronunciation, vocabulary studies, and listening comprehension, reflecting a more integrated model of spoken-language processing.

Another contribution lies in the examination of AI-generated speech, an emerging area of global interest. While international research has begun to explore the pedagogical potential of synthetic voices, few studies have systematically documented the

error patterns in AI narration or their implications for learners. The findings here provide early evidence that AI speech, despite its fluency, introduces predictable segmental and suprasegmental inaccuracies that can mislead learners—an insight that positions this work at the forefront of technology-mediated pronunciation research.

Finally, the author's long-term focus on interpreting, decoding, and auditory discrimination adds a cognitive dimension that complements global models of pronunciation as a perceptual–motor skill. By demonstrating how decoding efficiency, lexical retrieval, and contextual knowledge interact with phonological processing, this research advances a more comprehensive understanding of pronunciation as a multi-component cognitive activity rather than a purely articulatory one.

Taken together, these contributions position the author's research program as a significant and distinctive voice within global pronunciation scholarship, one that bridges phonology, cognition, technology, and applied linguistics, and enriches the international understanding of how learners perceive, interpret, and produce spoken language.

4.7 *Comparison With Previous Systematic Reviews and Meta-Analyses in the Literature*

When compared with the growing body of SRs and MAs in L2 pronunciation research, the present SR occupies a distinct position in both scope and focus. Existing reviews tend to cluster around several dominant themes: (1) instructional effectiveness, including explicit phonetic instruction, articulatory training, shadowing, and high-variability phonetic training (Chau & Huensch, 2025; Yağız et al., 2024; Mahdi & Mohsen, 2024; Whitworth, 2024; Vasquez & Aguirre, 2025); (2) technology-mediated pronunciation learning, particularly mobile-assisted learning, apps, podcasts, and AI-based tools (Metruk, 2024; Stoughton & Kang, 2024; Aryanti & Santosa, 2024; Muslimah & Tarihoran, 2025; Meidyana et al., 2025); (3) context-specific reviews, such as pronunciation teaching in CFL contexts, Malaysian EFL settings, Bangladeshi CLT classrooms, and Arabic phonetic instruction for Hijaiyyah (Li et al., 2021; Hamzah et al., 2015; Qader et al., 2024; Mohamed et al., 2024); and (4) specialized subdomains, including intelligibility and accentedness, poetic style, grapheme-to-phoneme modelling, pronunciation for speech-impaired children, and NNEST beliefs (Ariningsih, 2023; Darejeh et al., 2019; Irie et al., 2023; Wang et al., 2024). Across these reviews, the primary emphasis is on instructional interventions, technology, teacher knowledge, or intelligibility outcomes. Very few reviews examine pronunciation as an integrated cognitive–linguistic process involving auditory discrimination, decoding, lexical access, background knowledge, and real-time meaning construction. None of the existing reviews synthesize research on interpreting, AI-generated speech errors, segmental error typologies among Arab learners, or the interaction between phonological transfer and domain-specific terminology—all of which are central to the present review.

Moreover, while several reviews address Arabic phonetics or pronunciation challenges in specific Arab contexts (Mohamed et al., 2024; Alrickaby, 2025), they focus primarily on Arabic sound systems, Hijaiyyah acquisition, or general EFL vowel difficulties, rather than the complex cross-linguistic mechanisms documented in this review, such as gemination transfer, faulty segmentation of long words, mispronunciation of foreign Proper Nouns, and the cognitive load of interpreting unfamiliar media discourse.

Finally, although recent reviews have begun to explore AI-mediated pronunciation learning, none has systematically analyzed AI-generated pronunciation errors or their pedagogical risks. This positions the present review at the forefront of emerging research on the reliability of synthetic speech as instructional input.

In sum, while previous SRs have advanced the field in areas such as instructional methods, technology integration, and intelligibility research, the present SR contributes a **unique, multi-layered synthesis** that integrates phonology, cognition, decoding, background knowledge, and AI-mediated input, offering a perspective not yet represented in the global pronunciation literature.

4.8 *Limitations of The Systematic Review*

Although this meta-analysis synthesizes a substantial body of research on pronunciation instruction and practice, the studies included do not represent all possible instructional techniques, practice activities, or technological tools used in L2 pronunciation teaching. The review reflects the specific methods examined across the author's research program, which emphasizes decoding-based instruction, guided oral reading, listening-supported practice, and selected forms of technology such as text-to-speech systems, mobile lessons, and AI-generated speech. Other instructional approaches—such as communicative pronunciation tasks, drama-based training, suprasegmental-focused interventions, or advanced speech-analysis software—fall outside the scope of the included studies. Similarly, the technological tools represented in the corpus do not encompass the full range of CAPT, ASR, or MALL applications currently available. As a result, the findings offer a coherent but a selective view of pronunciation instruction, shaped by few instructional and technological domains investigated in the reviewed studies.

4.9 *Future Research Directions*

The synthesis of findings across the five clusters highlights several suggestions for future research. First, future research should investigate longitudinal developmental studies, tracing how phonological transfer, decoding skills, and lexical access evolve over time. Longitudinal designs would clarify whether certain error types (e.g., gemination, faulty segmentation, mispronunciation of Proper Nouns) diminish with exposure or persist as fossilized patterns. There is a need for experimental intervention studies that integrate pronunciation training with decoding instruction, vocabulary development, and background-knowledge building. Most existing interventions focus on isolated phonetic training. Future work should test integrated models that reflect the multi-component nature of pronunciation revealed in this review. The inaccuracies in AI-generated speech call for systematic evaluations of different AI engines, languages, and voice models. Future studies may examine how learners perceive, internalize, and reproduce AI-generated errors, and whether certain AI voices are more reliable than others for pronunciation instruction, especially in diacritic-sensitive languages like Arabic.

Additionally, interpreting research would benefit from cognitive-load studies that use eye-tracking, reaction-time measures, or neurocognitive methods to explore how learners process unfamiliar Proper Nouns, acronyms, and technical terminology in real time. Such methods could illuminate the moment-to-moment breakdowns that lead to nonsense forms, rhyming approximations, or literal translations.

Moreover, future research should explore domain-specific pronunciation challenges in professional fields such as healthcare, diplomacy, media, and aviation. These contexts involve dense terminology and high-stakes communication, yet remain underexplored in pronunciation research.

Finally, there is a need for corpus-based and computational approaches that model pronunciation errors among Arabic speakers, including grapheme-to-phoneme mapping, error prediction, and automated feedback systems. Such tools could support both research and pedagogy by identifying recurrent patterns and generating targeted instructional materials.

Together, these directions point toward a research agenda that is broader, more interdisciplinary, and more technologically informed—one that builds on the foundations established in this review and advances our understanding of pronunciation as a complex cognitive-linguistic skill.

5. Recommendations

To improve pronunciation accuracy, auditory discrimination, and decoding skills among Arabic–English learners, and enhance the reliability of instructional input, the current study recommends that instruction should combine phonological training with decoding practice and systematic vocabulary development, especially for Proper Nouns, technical terminology, and domain-specific expressions frequently encountered in media, healthcare, and academic contexts. Learners benefit from targeted instruction on features strongly influenced by Arabic phonotactics, including consonant clusters, stress placement, vowel quality, and the avoidance of over-gemination. Raising students' awareness of cross-linguistic contrasts can help prevent fossilization of L1-based patterns. Since many pronunciation and meaning-transfer errors stem from unfamiliarity with global institutions, geographical names, acronyms, and political designations, curricula should incorporate structured exposure to high-frequency international terminology and real-world media content. Educators and learners should critically evaluate synthetic audio before using it as a pronunciation model. When possible, AI input should be supplemented with human-produced recordings, especially for diacritic-sensitive Arabic words and context-dependent English forms. Furthermore, learners need systematic practice in distinguishing unfamiliar phonemes, decoding long or complex words, and recognizing Proper Nouns in fast speech. Activities such as minimal-pair training, slowed-speech exposure, and repeated listening to authentic media can strengthen these skills. Instruction should include timed listening, shadowing, and short interpreting segments that simulate real-world processing demands. These tasks help learners develop rapid form–meaning mapping and reduce reliance on compensatory strategies. Specialized fields such as healthcare, diplomacy, and media require tailored pronunciation lists, annotated corpora, and practice materials that reflect the terminology learners encounter in professional settings. Learners should be trained to recognize when they are relying on approximation strategies, such as nonsense forms, rhyming analogies, or literal translation, and to replace these with more accurate decoding and inference strategies.

6. conclusion

This SR brings together two decades of research that collectively illuminate the complex, multi-layered nature of pronunciation performance in Arabic–English contexts. Across the five thematic clusters, AI-generated speech, segmental and suprasegmental errors, decoding and auditory discrimination, interpreting accuracy, and domain-specific pronunciation, the evidence converges on a central insight: pronunciation is not an isolated phonetic skill but a cognitive–linguistic process shaped by phonological transfer, lexical access, decoding efficiency, background knowledge, and input quality. The studies reviewed here demonstrate that when any of these components is weak, learners resort to predictable compensatory strategies such as nonsense forms,

rhyming analogies, literal translation, or overgeneralization, revealing the cognitive pressures involved in real-time speech processing. The review also highlights the emerging role of AI-generated speech as both a resource and a risk. While synthetic voices offer accessible models for learners, they introduce systematic pronunciation errors—particularly in diacritic-sensitive Arabic words and context-dependent English forms—that can mislead learners and reinforce inaccurate patterns. This underscores the need for critical evaluation of AI tools and for pedagogical frameworks that integrate human-produced input with technology-mediated resources.

By synthesizing findings from pronunciation, listening comprehension, decoding, interpreting, and professional communication, this review positions pronunciation as an interdisciplinary domain that intersects with cognition, vocabulary studies, and applied linguistics. It also fills a gap in global pronunciation research by offering a detailed account of the unique challenges faced by Arabic speakers—challenges that are often overlooked in broader multilingual reviews.

Ultimately, the review calls for instructional approaches and research agendas that reflect the integrated nature of pronunciation, combining phonological training with decoding practice, vocabulary expansion, background-knowledge building, and careful selection of input sources. In doing so, it lays the groundwork for more comprehensive, context-sensitive, and technologically informed models of pronunciation teaching and learning—models that better reflect the realities of learners navigating between Arabic and English in academic, professional, and digital environments.

Conflicts of Interest: The author declares no conflict of interest.
ORCID ID: https://orcid.org/0000-0002-6255-1305

References

[1] Al-Jarf, R. (2026a). A self systematic review of translation error studies (2000–2025): The Case of students' errors in English–Arabic and Arabic–English translation. *International Journal of Translation and Interpretation Studies*, *6*(1), 16-32. DOI: 10.32996/ijtis.2026.6.1.2. Google Scholar

[2] Al-Jarf, R. (2026b). Arabic–English Transliteration of Personal Names and Public Signages: A Systematic Review and Meta-analysis. *British Journal of Applied Linguistics*, *6*(1), 01-14. DOI: 10.32996/bjal.2025.6.1.1. Google Scholar

[3] Al-Jarf, R. (2026c). Children's Language Acquisition and Development in Saudi Arabia: A Systematic Review and Meta Analysis. *Journal of Learning and Development Studies*, *6*(1), 18-37. DOI: 10.32996/jlds.2026.6.1.3. Google Scholar

[4] Al-Jarf, R. (2026d). Innovative Word Formation and Pluralization Processes in Arabic: A Systematic Review. *Journal of Humanities and Social Sciences Studies*, *8*(1), 44-60. DOI: 10.32996/jhsss.2026.8.1.6. Google Scholar

[5] Al-Jarf, R. (2026e). Systematic review and meta-analysis of 2024–2025 studies on AI Arabic translation, linguistics and pedagogy. *Frontiers in Computer Science and Artificial Intelligence*, *5*(1), 07-27. DOI: 10.32996/jcsts.2026.5.1.2. Google Scholar

[6] Al-Jarf, R. (2025a). Arabic transliteration of borrowed english nouns with /g/ by Artificial Intelligence (AI). *Journal of Computer Science and Technology Studies*, *7*(9), 245-252. https://doi.org/10.32996/jcsts.2025.7.9.29. Google Scholar

[7] Al-Jarf, R. (2025b). *Can students learning arabic as a foreign language use Arabic YouTube videos narrated by Artificial Intelligence (AI) for Listening Practice*. 2nd International Forum on Teaching Arabic in the Modern World: Traditions and Innovations. Sheikha Fatima bint Mubarak Center for Education. Primakov International School Moscow, Russia.November 15–16, 2025. Google Scholar

[8] Al-Jarf, R. (2025c). Faulty consonant gemination in the pronunciation of English biomedical terms by Arab healthcare professionals. *Journal of Medical and Health Studies, 6*(3), 56-66. DOI: 10.32996/jmhs.2025.6.3.9. Google Scholar

[9] Al-Jarf, R. (2024d). Intonational Meanings of Discourse Markers in Spoken Colloquial Arabic. Journal of Pragmatics and Discourse Analysis, 3(2), 01-10. Google Scholar

[10] Al-Jarf, R. (2025e). Mapping pronunciation errors in English silent consonants: A corpus-based study of Saudi EFL undergraduates. *Journal of Humanities and Social Sciences Studies, 7*(6), 13-21. DOI: 10.32996/jhsss.2025.7.6.2. Google Scholar

[11] Al-Jarf, R. (2025f). *Pronunciation errors in AI-narrated Arabic YouTube videos*. LICCS Pronunciation Errors in Arabic YouTube Videos Narrated Online Conference on Teaching and Research in Language and Culture: Past, Present and AI. Babeș-Bolyai University, Cluj-Napoca, Romania. September 11-12, 2025. Google Scholar

[12] Al-Jarf, R. (2025g). Pronunciation errors in Arabic YouTube videos narrated by AI. *Frontiers in Computer Science and Artificial Intelligence*, *4*(2), 01-12. https://doi.org/10.32996/fcsai.2025.2.2.1. Google Scholar

[13] Al-Jarf, R. (2025h). Splitting unsplittable foreign words in casual speech by EFL Arab learners. *British Journal of Applied Linguistics, 5*(2), 01-11. https://doi.org/10.32996/bjal.2025.5.2.1. Google Scholar

[14] Al-Jarf, R. (2025i). Vowel pronunciation errors in English biomedical terminology by Arab healthcare professionals. *Journal of Medical and Health Studies, 6*(2), 145-155. https://doi.org/10.32996/jmhs.2025.6.2.22. Google Scholar

[15] Al-Jarf, R. (2023a). Absence of vowels in the English spelling of Arabic personal names on social media. *International Journal of English Language Studies (IJELS), 5*(4), 88-97. Doi: 10.32996/ijels.2023.5.4.7. ERIC ED633828. Googe Scholar

[16] Al-Jarf, R. (2023b). Ambiguity in Arabic negative polar questions. *Journal of Pragmatics and Discourse Analysis, 2*(1),47-58. DOI: 10.32996/jpda.2023.2.1.6 ERIC ED628264. Google Scholar

[17] Al-Jarf, R. (2023c). Clipping of borrowings in spoken Arabic. *International Journal of Linguistics, Literature and Translation, 6*(1), 68-76. ERIC ED633842 https://doi.org/10.32996/ijllt.2023.6.11.9. Google Scholar

[18] Al-Jarf, R. (2023d). English spelling of Arabic compound personal names by educated Arabs on Facebook. *Journal of Humanities and Social Sciences Studies (JHSCS), 5*(1), 53-64. DOI: 10.32996/jhsss.2023.5.1.8. Google Scholar

[19] Al-Jarf, R. (2023e). Grammar podcasts for ESL college students in distance learning. *British Journal of Teacher Education and Pedagogy, 2*(2), 36-42. https://doi.org/10.32996/bjtep.2023.2.2.4. ERIC ED628487. Google Scholar

[20] Al-Jarf, R. (2023f). English spelling of the glottal stop and voiced pharyngeal fricative in Arabic personal names by educated Arabs on Facebook. *International Journal of English Language Studies, 5*(1), 11-22. DOI: 10.32996/ijels.2023.5.1.2. Google Scholar

[21] Al-Jarf, R. (2023g). Semantic and syntactic anomalies of Arabic-transliterated compound shop names in Saudi Arabia. *International Journal of Arts and Humanities Studies (IJAHS), 3*(1), 1-8. DOI: 10.32996/Ijahs.2023.3.1.1. Google Scholar

[22] Al-Jarf, R. (2023h). Testing multiple vocabulary associations for effective long-term learning. *British Journal of Teacher Education and Pedagogy, 2*(3), 57-71. DOI: 10.32996/bjtep.2023.2.3.6. ERIC ED634388. Google Scholar

[23] Al-Jarf, R. (2022a). A multiple-associations approach to teaching technical terms in English for specific purposes courses A multiple-associations approach to teaching technical terms in English for specific purposes courses. *International Journal of English Language Studies (IJELS), 4*(2), 56-66. DOI: 10.32996/ijels.2022.4.2.5. ERIC ED621773. Google Scholar

[24] Al-Jarf, R. (2022b). Deviant Arabic transliterations of foreign shop names in Saudi Arabia and decoding problems among shoppers. *International Journal of Asian and African Studies (IJAAS), 1*(1), 17-30. DOI: 10.32996/ijaas.2022.1.1.3. Google Scholar

[25] Al-Jarf, R. (2022c). Developing Students' Global Awareness in EFL Reading and Speaking. *South Asian Research Journal of Arts, Language and Literature, 4*(1), 31-38. DOI: 10.36346/sarjall.2022.v04i01.005. **ERIC** ED618471. Google Scholar

[26] Al-Jarf, R. (2022d). English-Arabic and Arabic-English interpreting competence of undergraduate student interpreters: A comparative study of directionality. *International Journal of Social Science and Education Research Studies (ijssers), 2*(1), 39–44. ERIC ED618628. Google Scholar

[27] Al-Jarf, R. (2022e). English transliteration of Arabic personal names with the definite article /al/ on Facebook. *British Journal of Applied Linguistics (BJAL), 2*(2), 23-37. DOI: 10.31926/but.pcs.2022.64.15.2.2. Google Scholar

[28] Al-Jarf, R. (2022f). Gemination errors in Arabic-English transliteration of personal names on Facebook. *International Journal of Linguistics Studies (IJLS), 2*(2),163-170. DOI: 10.32996/ijls.2022.2.2.18. Google Scholar

[29] Al-Jarf, R. (2022g). *Integrating current global events and technology in interpreting practice*. International Innovative Pedagogical Technologies in Teaching Foreign Languages. Perm State University, Russia. April 4. Google Scholar

[30] Al-Jarf, R. (2022h). Proper noun pronunciation inaccuracies in English by Educated Arabic speakers. *British Journal of Applied Linguistics (BJAL), 4*(1), 14-21. https://doi.org/10.32996/bjal.2022.2.1.3. **ERIC** ED619388. Google Scholar

[31] Al-Jarf, R. (2022i). Student-interpreters' foreign proper noun pronunciation errors in English-Arabic and Arabic-English media discourse interpreting. *International Journal of Translation and Interpretation Studies (IJTIS), 2*(1), 80-90. Doi: 10.32996%2Fijtis.2022.2.1.11. ERIC ED619940. Google Scholar

[32] Al-Jarf, R. (2022j). Text-to-speech software as a resource for independent interpreting practice by undergraduate interpreting students. *International Journal of Translation and Interpretation Studies (IJTIS), 2*(2), 32-39. DOI: 10.32996/ijtis.2022.2.2.3. ERIC ED621859. Google Scholar

[33] Al-Jarf, R. (2022k). Text-to-speech software for promoting EFL freshman students' decoding skills and pronunciation accuracy. *Journal of Computer Science and Technology Studies (JCSTS), 4*(2), 19-30. DOI: 10.32996/jcsts.2022.4.2.4. ERIC ED621861. Google Scholar

[34] Al-Jarf, R. (2022l). YouTube videos as a resource for self-regulated pronunciation practice in EFL distance learning environments. *Journal of English Language Teaching and Applied Linguistics, 4*(2), 44-52. DOI: 10.32996/jeltal.2022.4.2.4. ERIC ED618965. Google Scholar

[35] Al-Jarf, R. (2022m). Variant transliterations of the same Arabic personal names on Facebook. *International Journal of English Language Studies (IJELS), 4*(4), 79-90. DOI: 10.32996/ijels.2022.4.4.11. Google Scholar

[36] Al-Jarf, R. (2021a). EFL female college students and instructors' preferred method of speaking assessment: A perspective from Saudi Arabia. Asian Journal of Education and Social Studies (AJESS), 16 (3), 38-50. doi: 10.9734/ajess/2021/v16i330403. ERIC ED616903. Google Scholar

[37] Al-Jarf, R. (2021b). EFl speaking practice in distance learning during the coronavirus pandemic 2020-2021. *International Journal of Research - GRANTHAALAYAH, 9*(7), 179-196. doi: 10.29121/granthaalayah.v9.i7.2021.4094. ERIC ED615084. Google Scholar

[38] Al-Jarf, R. (2021c). Feasibility of digital multimedia language labs for interpreting instruction as perceived by interpreting instructors in Saudi Arabia. *International Journal of Linguistics, Literature and Translation (IJLLT), 4*(4), 70-79. Doi: 10.32996/ijllt.2021.4.4.8. ERIC ED613819. Google Scholar

[39] Al-Jarf, R. (2021d). Integrating participation goals in writing activities for EFL college students. *Journal for Research Scholars and Professionals of English Language Teaching (JRSP-ELT), 23*(5), January. ERIC ED613141. Google Scholar

[40] Al-Jarf, R. (2021e). Mobile audiobooks, listening comprehension and EFL college students. *International Journal of Research – GRANTHAALAYAH, April 9*(4), 410-423. DOI: 10.29121/granthaalayah.v9.i4.2021.3868 Google Scholar

[41] Al-Jarf, R. (2021f). Teaching interpreting for tourism purposes. *Journal of English Language Teaching and Applied Linguistics (JELTAL), 3*(13), 17-26. https://doi.org/10.32996/jeltal.2021.3.13.3. ERIC ED618208. Google Scholar

[42] Al-Jarf, R. (2021g). TED talks as a listening resource in EFL college classrooms. *International Journal of Language and Literary Studies (IJLLS), 2*(3), 256–267. DOI: 10.36892/ijlls.v2i3.691. ERIC ED615127. Google Scholar

[43] Al-Jarf, R. (2020a). *Integrating TED lectures in EFL college listening practice*. 25th TCC worldwide Online Conference. April 14-16, 2020. https://www.researchgate.net/profile/Reima-Al-Jarf/publication/356815037. Google Scholar

[44] Al-Jarf, R. (2020b). Mobile apps in the EFL College Classroom. *Journal for Research Scholars and Professionals of English Language Teaching (JRSP-ELT), 4*(22),1-5. (ISSN: 2456-8104). ERIC ED613138. Google Scholar

[45] Al-Jarf, R. (2019). *EFL freshman students' difficulties with phoneme-grapheme relationships*. 5th VietTESOL International Convention. Hue University of Foreign Languages, Hue, Vietnam. October 11-12. Google Scholar

[46] Al-Jarf, R. (2018a). *Effect of background knowledge on auditory comprehension in interpreting courses*. 4th Brno Conference on Linguistics Studies in English titled "Interpretation of Meaning in Spoken and Written Discourse. Masaryk University, Brno, Czech Republic. Sept 16-17. ERIC ED665097. Google Scholar

[47] Al-Jarf, R. (2018b). *Effect of background knowledge on auditory comprehension in interpreting courses*. In Renata Jancarikova (Ed.) Interpretation of Meaning across Discourse, pp. 97-108. Muni Press, Brno, Czech Republic. https://doi.org/10.5817/CZ.MUNI.M210-8947-2018. Google Scholar

[48] Al-Jarf, R. (2015a). *Enhancing reading and speaking skills in EFL through multicultural children's short stories*. 7th International Conference, Building Cultural Bridges (ICBCB) titled: Integrating Languages, Linguistics, Literature, Translation, and Journalism into Education.Suleyman Demire lUniversity, Almaty, Kazakhstan. April 23-24, 2015. https://doi.org/10.2139/ssrn.3848464. **ERIC ED**610158. Google Scholar

[49] Al-Jarf, R. (2015b). *What instructors and students should know about interpreting problems*. College of Languages and Translation, King Saud University, Riyadh, Saudi Arabia. January 28. Google Scholar

[50] Al-Jarf, R. (2012). *Mobile technology and student autonomy in oral skill acquisition*. In Javier E. Díaz Vera's Left to My Own Devices: Innovation and Leadership in English Language Teaching. 105-129. Brill. DOI: 10.1163/9781780526478_007. Google Scholar

[51] Al-Jarf, R. (2011). Teaching spelling with mind-mapping software. *Asian EFL Journal, 53*(July), 4-16. Google Scholar

[52] Al-Jarf, R. (2010). Spelling error corpora in EFL. *US-China Foreign Language. Sino-US English Teaching, 7*(1), 6-15. ERIC ED620777. Google Scholar

[53] Al-Jarf, R. (2009). *Auditory and visual problems of good and poor EFL college spellers.* College of Languages and Translation. https://www.researchgate.net/publication/238599084. Google Scholar

[54] Al-Jarf, R. (2008a). *Integrating global themes in EFL speaking instruction*. 2nd International Conference Language, Literature and Translation in an Interdependent World. University of Bahrain, Bahrain. April 8-10. Google Scholar

[55] Al-Jarf, R. (2008b). *Listening-spelling strategies in EFL Arab college students*. College of Languages of Translation, King Saud University Seminars. Google Scholar

[56] Al-Jarf, R. (2008c). *Phonological and orthographic problems in EFL college spelling*. First Regional Conference on English Language Teaching and Literature (ELTL 1). Islamic Azad University-Roudehen. TELLIS Conference Proceedings. ERIC ED611115. Google Scholar

[57] A-Jarf, R. (2008d). *Sources of spelling errors in EFL Arab college students*. College of Languages of Translation seminars, King Saud University. https://www.researchgate.net/profile/R. -Al-Jarf/publication/345900801 Google Scholar

[58] Al-Jarf, R. (2007a). *Faulty strategies of EFL freshman spellers, Saudi Arabia*. College of language and translation. King Saud University, Riyadh, Saudi Arabia. Google Scholar

[59] Al-Jarf, R. (2007b). *From reticence to fluency: The effect of TBLT on students' speaking ability*. The International Conference on Task-Based Language Teaching. Centre for Language and Migration, University of Leuven, Belgium. Google Scholar

[60] Al-Jarf, R. (2007c). *How to teach liaison interpreting to beginners*. Foundations for a Pedagogy of Arabic Translation Conference. University College Francisco Ferrer Brussels, Belgium. October 18-20. ERIC ED613512. Google Scholar

[61] Al-Jarf, R. (2006). *Making connections in vocabulary instruction*. 2nd ClaSic Conference proceedings. Singapore. ERIC ED497939. Google Scholar

[62] Al-Jarf, R. (2005a). *The effects of listening comprehension and decoding skills on spelling achievement of EFL Freshman students*. *English language and literature Education*. *Journal of the English Language Teachers in Korea (ETAK), 11*(2). ERIC ED625524. Google Scholar

[63] Al-Jarf, R. (2005b). The relationship among spelling, listening and decoding skills in EFL freshman students. *English Language & Literature Teaching, 11(2), 35-55*. Google Scholar

[64] Al-Jarf, R. (2000). *Bridging the gap between teacher and learner in liaison interpreting*. TESOL Arabia' 2000 Conference, "Bridging the Gap between Teacher and Learner". Al-Ain, UAE, April 12-14. Google Scholar

[65] Al-Jarf, R. (1999). *Listening-spelling strategies of freshmen students*. TESOL Arabia Conference titled "Unity and diversity. Google Scholar

[66] Ahn, Y. (2013). *Pronunciation instruction in English as a foreign language contexts: A review of goals and best practices*. M.A. Thesis. The University of Texas at Austin

[67] Almelhes, S. (2023). A review of artificial intelligence adoption in second-language learning. *Theory and Practice in Language Studies*, *13*(5), 1259-1269.

[68] Almusharraf, A., et al. (2024). The effect of settings, educational level and tools on computer-assisted pronunciation training: A meta-analysis. *Journal of Computer Assisted Learning*, *40*(4), 1605-1615.

[69] Alrickaby, A. (2025). Research on Iraqi EFL Learners' Pronunciation: A Review Article. *Journal of World Englishes and Educational Practices*, *7*(1), 32-36.

[70] Amrate, M., & Tsai, P. (2024). Computer-assisted pronunciation training: A systematic review. *ReCALL*, 1-21.

[71] Ariningsih, T. (2023). Impact of poetic style for pronunciation: A systematic literature review. *INJURITY: Journal of Interdisciplinary Studies*, *2*(5), 390-400.

[72] Aryanti, R. & Santosa, M. (2024). A systematic review on artificial intelligence applications for enhancing EFL students' pronunciation skill. *The Art of Teaching English as a Foreign Language (TATEFL)*, *5*(1), 102-113.

[73] Chau, T., & Huensch, A. (2025). The relationships among L2 fluency, intelligibility, comprehensibility, and accentedness: A meta-analysis. *Studies in Second Language Acquisition*, *47*(1), 282-307.

[74] Darejeh, A., Salim, S. & Asemi, A. (2019). Speech pronunciation practice system for speech-impaired children: a systematic review of impacts and functionality. *Universal Access in the Information Society*, *18*(1), 169-189.

[75] Dhouib, A., et al. (2022). Arabic automatic speech recognition: A systematic literature review. *Applied Sciences*, *12*(17), 8898. https://doi.org/10.3390/app12178898.

[76] Hamzah, M., Norain, N. & Abdullah, A. (2015). Pronunciation research in the Malaysian context: A systematic review. *LSP International Journal*, *2*(1), 63–73.

[77] Irie, E., Juan, S. & Saee, S. (2023). A review on grapheme-to-phoneme modelling techniques to transcribe pronunciation variants for under-resourced language. *Pertanika Journal of Science & Technology*, *31*(3).

[78] Jeong, H., & Lindemann, S. (2025). Beyond ideologies of nativeness in the intelligibility principle for L2 English pronunciation: A corpus-supported review. *System, 103599*. DOI: 10.1016/j.system.2025.103599

[79] Jie, M. (2022). *Computer assisted pronunciation training (CAPT): A systematic review of studies from 2012 to 2021*. In International Conference on Computers in Education (pp. 575-580).
[80] Kaur, S., & Bhathal, G. (2026). Exploring speech recognition and pronunciation accuracy detection techniques: A comprehensive review. *Recent Advances in Computational Methods in Science and Technology*, 76-84.
[81] Kunova, R. (2025) Foundations and future directions in pronunciation teaching: a literature review. *Slavonic Pedagogical Studies Journal, 14*(1). DOI: 10.18355/PG.2025.14.1.1
[82] Li, L. (2018). A literature review of research on pronunciation instructional strategies in Chinese as a foreign language (CFL). *EDULEARN18 Proceedings*, 4935-4939.
[83] Li, L., et al. (2021). Pedagogical content knowledge: A systematic review of Chinese language pronunciation teaching in the CFL context. *Foreign Language Annals*, *54*(2), 525-557.
[84] Mahdi, H. & Al Khateeb, A. (2019). The effectiveness of computer-assisted pronunciation training: A meta-analysis. *Review of Education*, *7*(3), 733-753.
[85] Mahdi, H. & Mohsen, M. (2024). Enhancing pronunciation learning through high variability phonetic training: A meta-analysis. *Language Teaching Research Quarterly*, *40*, 29-45.
[86] Meidyana, E., Salsabila, F. & Harris, S. (2025). AI for all voices: A scoping review on inclusive pronunciation and speaking AI tools. *English Language and Literature International Conference (ELLiC),* Vol. 8, pp. 428-441.
[87] Metruk, R. (2024). Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, *29*(13), 16255-16282.
[88] Mohamed, Y., et al. (2024). A systematic review of arabic phonetic: Hijaiyyah's pronunciation among new learners. *Ijaz Arabi Journal of Arabic Learning*, *7*(2).
[89] Muslimah, N., & Tarihoran, N. (2025). MALLing your way to perfect pronunciation: A systematic review (2020-2024). *INJOSEDU: International Journal of Social and Education*, *2*(6), 1702-1715.
[90] Ngo, T., Chen, H., & Lai, K. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, *36*(1), 4-21.
[91] Qader, M. et al. (2024). Challenges of pronunciation practices in the ESL curriculum within the CLT framework in Bangladesh: A systematic review. *World Journal of Education*, *14*(3), 137-147.
[92] Ramirez, M. (2024). The use of podcasts for language learning: A systematic review of literature. *International Journal of English Language Studies*, *6*(1), 47-53.
[93] Stoughton, A. & Kang, O. (2024). A systematic review of empirical mobile-assisted pronunciation studies through a perception–production lens. *Languages*, *9*(7), 251.
[94] Suseno, E. (2023). Developing pronunciation skills through application to enhance speaking ability: Systematic literature review. *Innovare Journal of Social Sciences*, *11*(5), 1-5.
[95] Terrier, L., Garnier, M., & Ali, S. (2025). Leading a scoping review on L2 pronunciation: Some key elements of methodology. *Research Methods in Applied Linguistics*, *4*(3), 100216.
[96] Ting, X., Kadir, Z. & Husain, S. (2025). Technology-integrated blended teaching and its impact on English pronunciation accuracy among EFL learners in China: A Bibliometric and Systematic Review. *Journal of Advanced Research in Social and Behavioural Sciences*, *41*(1), 1-14.
[97] Tseng, W., et al. (2022). The effects of MALL on L2 pronunciation learning: A meta-analysis. *Journal of Educational Computing Research*, *60*(5), 1220-1252.
[98] Vasquez, E. & Aguirre, S. (2025). The efficacy of explicit articulatory phonetics instruction in L2 pronunciation acquisition: A systematic review of pedagogical approaches for Spanish L1 Learners of English. *Boletín Científico Ideas y Voces*, *5*(3), ág-427.
[99] Wahyudi, D., et al. (2025). Improving students' pronunciation using word up application: a systematic literature review. *Jurnal Intelek Insan Cendikia*, *2*(12), 18735-18739.
[100] Wang, J., et al. (2024). Voices unveiled: Navigating NNESTs' beliefs on pronunciation teaching-a systematic review. *Forum for Linguistic Studies, 6*(6), pp. 915-928. Whioce Publishing Pte. Ltd.
[101] Whitworth, B. (2024). *Shadowing for pronunciation: a systematic review*. University of Oxford
[102] Whitworth, B., & Rose, H. (2025). A systematic review of research on the use of shadowing for second language pronunciation teaching. *Research Synthesis in Applied Linguistics*, 1-31.
[103] Yağız, O., Kaya, F., & Ötügen, R. (2024). The effectiveness of L2 pronunciation instruction: A critical systematic review of the intervention studies. *The Literacy Trek*, *10*(1), 21-41.