ARCHAIA JOURNAL OF SCIENCE AND TECHNOLOGY
VOL. 1, NO. 3, Sep 2024
https://doi.org/10.48161/ajst.v1n3a86

The Effect of Mother Tongue Interference on English Pronunciation among EFL Learners

Djuwari Djuwari[1*], Vinoth Kumar Chokkalingam[2]

[1] English Department, Universitas Nahdlatul Ulama Surabaya (UNUSA), Surabaya, East Java, Indonesia;
[2] Department of English Linguistics, Navoi State University, Navoi, Uzbekistan;

* **Corresponding author:** djuwarisehat@gmail.com.

ABSTRACT: This study examines transfer-related first-language influence on English pronunciation among English as a Foreign Language learners and distinguishes communication-relevant pronunciation difficulty from legitimate accent variation. Using a mixed-methods controlled-task sociophonetic design, the study analyzed the speech of 180 university-level EFL learners from Arabic, Spanish, Chinese, and French first-language backgrounds. The study collected data through controlled word-list production, a reading passage, spontaneous speech tasks, pronunciation perception experiments, language-background questionnaires, listener ratings, acoustic analysis, and learner interviews. The quantitative strand used descriptive statistics, group comparisons, acoustic measures, regression analysis, and mixed-effects models, while the qualitative strand used thematic coding to interpret learner experiences of pronunciation anxiety, accent identity, self-monitoring, and classroom feedback. The results indicate that transfer-related patterns were most visible in vowel quality, consonant substitution, final-cluster simplification, lexical stress placement, and rhythm timing. L1 phonological distance predicted pronunciation-transfer difficulty, but English proficiency, exposure, phonetic awareness, pronunciation training, and frequency of communicative use moderated this relationship. Listener ratings confirmed that accentedness and intelligibility were related but not equivalent: several accented features remained comprehensible, whereas stress displacement and cluster reduction produced stronger processing costs. The study contributes an integrated sociophonetic pronunciation-transfer model linking cross-linguistic influence, acoustic evidence, listener perception, learner identity, and intelligibility-oriented pedagogy. The findings suggest that EFL pronunciation instruction should prioritize communication-relevant features while avoiding deficit-based views of multilingual accent variation.

Keywords: mother tongue interference; English pronunciation; EFL learners; phonological transfer; intelligibility.

I. INTRODUCTION

English pronunciation remains a major concern in English as a Foreign Language education because higher education, academic mobility, workplace communication, digital interaction, tourism, and intercultural participation increasingly require spoken English. Many EFL learners develop substantial grammatical and lexical knowledge but still struggle to make their spoken English easily understood. This difficulty is not merely technical. Pronunciation influences participation, confidence, perceived competence, and access to communicative opportunities. However, the present study approaches pronunciation difficulty cautiously: it does not treat accent variation as failure, but examines which transfer-related features create measurable processing costs for listeners and which features remain acceptable markers of multilingual speech.

In this study, mother tongue interference refers to cross-linguistic phonological influence rather than linguistic deficiency. Learners do not approach English sounds as neutral listeners; they perceive and produce English through sound categories, syllable structures, stress patterns, rhythm, and intonation resources shaped by previous language experience. Speech Learning Theory and Perceptual Assimilation explain that learners often filter second-language sounds through existing phonetic categories and may assimilate unfamiliar English contrasts to the closest available first-language pattern [1], [2]. This transfer can support learning when categories overlap, but it can create difficulty when English uses contrasts, clusters, stress placement, or reduction patterns absent from, or distributed differently in, the learner's L1.

Current pronunciation pedagogy has moved away from the unrealistic expectation that EFL learners should sound like native speakers. Intelligibility-oriented research argues that the instructional goal should be clear, comprehensible, and socially respectful communication rather than accent elimination [3]. This distinction is central to the present study. A learner may retain a recognizable Arabic, Spanish, Chinese, or French accent and still communicate effectively. At the same time, some transfer-related patterns may reduce intelligibility by changing lexical identity, obscuring grammatical endings, weakening word recognition, or disrupting stress and rhythm. The key issue, therefore, is not whether speech sounds native-like, but whether specific phonological features affect listener understanding, comprehensibility, and learner confidence.

Recent meta-analytic and bibliometric work shows that second-language pronunciation research has expanded from isolated error analysis to broader concerns about measurement, instruction, intelligibility, perception, individual differences, and classroom applications [4], [5]. Nevertheless, many classroom studies still discuss mother tongue influence in general terms without identifying which phonetic or prosodic features matter most, how listeners respond to those features, and how learners themselves interpret pronunciation feedback. Stronger EFL pronunciation research should combine acoustic evidence, phonetic coding, listener judgment, perception tasks, and learner accounts rather than relying only on teacher impressions or error lists.

The international relevance of this problem is clear. In higher education, EFL learners often need spoken English for presentations, research seminars, interviews, online collaboration, and academic exchange. In professional settings, pronunciation can influence public service interaction, customer communication, teamwork, and credibility. Learners' first languages provide different phonological resources and constraints, but group descriptions must be handled with care because dialect, schooling, exposure, proficiency, and individual experience shape pronunciation outcomes. The present study therefore interprets Arabic, Spanish, Chinese, and French L1 patterns as sample tendencies rather than fixed cultural or linguistic traits.

The study's originality lies in moving beyond a list of pronunciation errors. It connects L1-influenced pronunciation patterns to acoustic measures, listener ratings, intelligibility, comprehensibility, accentedness, learner confidence, and pedagogical response. It also treats pronunciation as socially situated: listeners evaluate learner speech, classroom input shapes it, identity and anxiety affect it, and access to phonetic training supports or constrains it. The study contributes to applied linguistics, sociophonetics, phonetics and phonology, second-language acquisition, and EFL pedagogy by asking not only what learners pronounce differently, but which differences matter for communication and how instruction can respond responsibly.

II. RESEARCH PROBLEM

The first problem is phonological. EFL learners often transfer first-language sound categories and syllable patterns into English. This may result in vowel merger, consonant substitution, final-consonant deletion, consonant-cluster simplification, epenthesis, or stress displacement. When English contrasts are absent or organized differently in the first language, learners may struggle to perceive and produce them consistently.

The second problem is communicative. L1-influenced pronunciation patterns do not all affect communication equally. Some accented features are noticeable but intelligible, whereas others reduce listener understanding. A research design that treats every non-native-like feature as equally problematic risks confusing accent variation with communication difficulty.

The third problem is pedagogical. Pronunciation teaching in many EFL classrooms remains irregular, textbook-dependent, or overly focused on repetition without systematic phonetic awareness. Learners may

not receive enough explicit support for articulatory control, listening discrimination, stress, rhythm, intonation, or self-monitoring.

The fourth problem is methodological. Previous studies of mother tongue influence often lack the combined evidence needed for rigorous interpretation: acoustic measurements, listener ratings, inter-rater reliability, perception-task data, and learner interpretation. Strong pronunciation research requires transparent recording procedures, phonetic coding, acoustic analysis, listener evaluation, and qualitative evidence of learner experience in the same design.

III. RESEARCH OBJECTIVES AND RESEARCH QUESTIONS

1. RESEARCH OBJECTIVES

This study examines how transfer-related first-language influence shapes English pronunciation among EFL learners and how these patterns relate to intelligibility, comprehensibility, accentedness, and oral communication. The specific objectives are:

- To identify the most frequent English pronunciation difficulties associated with first-language phonological transfer.
- To analyze how L1 background relates to English vowels, consonants, consonant clusters, and syllable structure.
- To examine transfer-related patterns in word stress, rhythm, intonation, and connected speech.
- To evaluate how pronunciation accuracy relates to listener judgments of intelligibility, comprehensibility, and accentedness.
- To test whether proficiency, exposure, phonetic awareness, pronunciation training, and frequency of English use moderate transfer-related pronunciation difficulty.
- To present an intelligibility-oriented pedagogical framework for EFL pronunciation instruction that respects accent variation while addressing communication-relevant difficulty.

2. RESEARCH QUESTIONS

- RQ1. What pronunciation difficulties occur most frequently among EFL learners from different first-language backgrounds?
- RQ2. How are transfer-related patterns observed in English segmental pronunciation, including vowels, consonants, consonant clusters, and syllable structure?
- RQ3. How does first-language transfer influence suprasegmental pronunciation, including stress, rhythm, intonation, and connected speech?
- RQ4. Which transfer-related pronunciation features most strongly affect intelligibility, comprehensibility, accentedness, and oral communication?
- RQ5. How do English proficiency, exposure, phonetic awareness, pronunciation training, and frequency of use moderate transfer-related pronunciation difficulty?
- RQ6. What pedagogical recommendations can be derived for pronunciation teaching in EFL classrooms?

IV. LITERATURE REVIEW

1. CROSS-LINGUISTIC INFLUENCE AND SECOND-LANGUAGE SPEECH LEARNING

Cross-linguistic influence in pronunciation stems from the interaction between existing first-language categories and new English phonetic targets. The revised Speech Learning Model explains that learners retain the capacity for phonetic learning, but category formation depends on the quality, quantity, and distribution of input [1]. Perceptual assimilation theory similarly shows that learners may map nonnative sounds onto existing L1 categories, making some English contrasts difficult to perceive before they are difficult to produce [2]. These perspectives explain why L1-influenced pronunciation patterns are systematic rather than random, but they should not be interpreted as fixed or irreversible.

The contemporary studies treat L1 transfer as one component in a wider system of learner experience. Age of learning, input, phonetic awareness, explicit instruction, attention, motivation, and the social value

of speech forms affect pronunciation development. This means learners from the same L1 background may share tendencies but still differ substantially in proficiency, exposure, and instructional history.

2. INTELLIGIBILITY, COMPREHENSIBILITY, AND ACCENTEDNESS

The intelligibility principle has reshaped pronunciation research and pedagogy. Rather than judging learner speech only by native-like accuracy, researchers distinguish intelligibility, how much of a message is understood; comprehensibility, how easy speech is to understand; and accentedness, how different speech sounds from a familiar target variety [3], [6]. This distinction matters because strongly accented speech can remain intelligible, while small pronunciation differences may create misunderstanding when they affect high-functional-load contrasts or stress patterns.

Listener-judgment research shows that comprehensibility and accentedness are related but partly independent constructs. Segmental errors, fluency breakdowns, lexical choice, grammar, and prosody can all influence listener response [7], [8]. Meta-analytic evidence further suggests that comprehensible speech is not identical to native-like speech [9]. Therefore, evaluate L1-influenced pronunciation by its communicative impact, not by accent stigma.

3. SEGMENTAL TRANSFER: VOWELS, CONSONANTS, AND CLUSTERS

Segmental transfer appears when learners substitute unfamiliar English sounds with similar L1 sounds, collapse English vowel contrasts, weaken final consonants, or adjust clusters to fit the first-language syllable template. Research on functional load shows that not all segmental contrasts deserve the same instructional priority [10]. A contrast that frequently distinguishes English words may matter more for intelligibility than a rare contrast with limited communicative consequence.

Segmental accuracy is also affected by task type. Learners may perform better in word lists than in spontaneous speech because controlled tasks allow attention to individual sounds, while free speech requires simultaneous control of vocabulary, grammar, fluency, and interaction. This task effect supports a mixed task battery rather than a single reading passage.

4. SUPRASEGMENTAL TRANSFER: STRESS, RHYTHM, AND INTONATION

Suprasegmental features are often more difficult to teach and measure than individual sounds, but they are central to intelligibility. Instructional studies show that work on word stress, rhythm, and intonation can improve comprehensibility, especially when learners are given explicit awareness and communicative practice [11], [12]. Exposure and repetition also influence the acquisition of spoken word forms and pronunciation memory [13].

Suprasegmental transfer may appear when learners use L1 timing patterns in English, assign stress according to first-language routines, reduce English weak forms insufficiently, or use intonation contours that listeners interpret differently. These patterns can be tied to identity and accent, but they may also affect listener processing when they disrupt word recognition or sentence focus.

5. INDIVIDUAL DIFFERENCES, CLASSROOM INSTRUCTION, AND SOCIOPHONETIC PERSPECTIVE

Research on individual differences shows that L2 speech learning is shaped by experiential, cognitive, and sociopsychological factors. Learners with greater English exposure, higher phonetic awareness, and stronger communicative use often show better pronunciation outcomes even when L1 distance is substantial [14]. In academic settings, listeners may prioritize comprehensibility more than a native-like accent when judging acceptability [15]. This reinforces the importance of sociophonetic analysis: speech is not simply produced; it is heard, evaluated, and socially interpreted.

Older but influential research on suprasegmental salience, pronunciation instruction, and teacher cognition remains important for understanding classroom practice [16-18]. Later studies and reviews indicate that pronunciation instruction can be effective, but results vary according to measurement construct, task type, duration, feedback quality, and whether instruction targets segmentals, suprasegmentals, or both [19-23]. Recent classroom intervention and teacher-cognition studies further show that pronunciation

outcomes depend on instructional focus, feedback quality, materials design, teacher modeling, and task-based practice [24], [25], [26], [27], [28].

6. CRITICAL SYNTHESIS OF ACOUSTIC, PERCEPTUAL, AND CLASSROOM EVIDENCE

A major limitation in the literature is that acoustic studies, listener-judgment studies, and classroom intervention studies often operate separately. Acoustic analysis provides precise evidence of vowel quality, VOT, duration, stress timing, and pitch movement, but acoustic deviation does not automatically indicate communication failure. Listener-judgment research captures intelligibility, comprehensibility, and accentedness, but listener familiarity and attitudes may influence it. Classroom intervention studies show that instruction can improve pronunciation, yet they often report outcomes without enough acoustic or perceptual triangulation. This study addresses this divide by integrating acoustic measurement, listener evaluation, learner background variables, and interview-based interpretation. The statistical strand is supported by established mixed-effects modeling procedures [29], [30], while the qualitative strand follows trustworthiness principles in thematic analysis [31], [32].

Table 1. Literature-review synthesis and study positioning.

Research strand	Key references	Main insight	Use in this study
L2 speech learning and perception	[1], [2]	L2 sounds are filtered through existing phonetic and perceptual categories.	Explains why L1 background shapes English sound production.
Intelligibility and accentedness	[3], [6]-[10]	Accentedness is related to but not identical with intelligibility and comprehensibility.	Supports communication-oriented interpretation.
Instruction and classroom effects	[4], [11]-[13], [18]-[28]	Pronunciation instruction works best when targets, feedback, and measurement are explicit.	Justifies controlled tasks and pedagogical implications.
Sociophonetics and individual differences	[14]-[17], [20]-[23]	Pronunciation is shaped by exposure, identity, listener judgment, and classroom context.	Connects acoustic patterns with interviews.
Analytical procedures	[29]-[32]	Mixed-effects modeling and thematic analysis strengthen transparency in quantitative and qualitative analyses.	Supports the statistical and qualitative procedures.

V. RESEARCH GAP AND STUDY CONTRIBUTION

1. RESEARCH GAP

The first gap is linguistic. Many studies identify English pronunciation errors, but fewer systematically link those errors to L1 phonological distance, acoustic deviation, and listener-rated intelligibility in the same design. Error lists are useful, but they do not explain which features have the greatest communicative effect.

The second gap is methodological. Some pronunciation studies rely solely on teacher judgment or reading-aloud tasks. Such designs may miss how learners sound in spontaneous speech and how listeners respond to that speech. A more rigorous approach needs controlled production, reading, spontaneous speech, perception experiments, acoustic analysis, listener ratings, questionnaires, and interviews.

The third gap is pedagogical and sociolinguistic. Researchers often frame L1-influenced pronunciation negatively, as if any accent is a problem. Research needs to distinguish accent variation from communication-relevant difficulty and connect findings to intelligibility-oriented instruction.

https://doi.org/10.48161/ajst.v1n3a86

2. STUDY CONTRIBUTION

The study contributes an integrated sociophonetic pronunciation-transfer model that links cross-linguistic phonological influence, acoustic evidence, listener evaluation, learner interpretation, and pedagogical response. The contribution is not simply to document pronunciation differences, but to identify which transfer-related features carry communicative weight and how EFL instruction can respond without treating multilingual accents as deficiencies.

VI. CONCEPTUAL AND THEORETICAL FRAMEWORK

The conceptual framework integrates Speech Learning Theory, Perceptual Assimilation, cross-linguistic influence, sociophonetics, phonetics and phonology, and intelligibility-oriented pedagogy. Theoretical constructs explain how L1 phonology shapes perception and production; measured variables include acoustic indices, phonetic coding, perception-task accuracy, listener ratings, learner-background variables, and interview themes; pedagogical outputs translate communication-relevant findings into instructional priorities.

The framework treats L1-influenced pronunciation as a patterned interaction between phonological transfer, sociophonetic experience, and communication. It does not assume that native-like pronunciation is the only desirable outcome. Instead, it asks which L1-influenced features affect intelligibility, which accent variations are acceptable, and how instruction can target the features that matter most for communication.

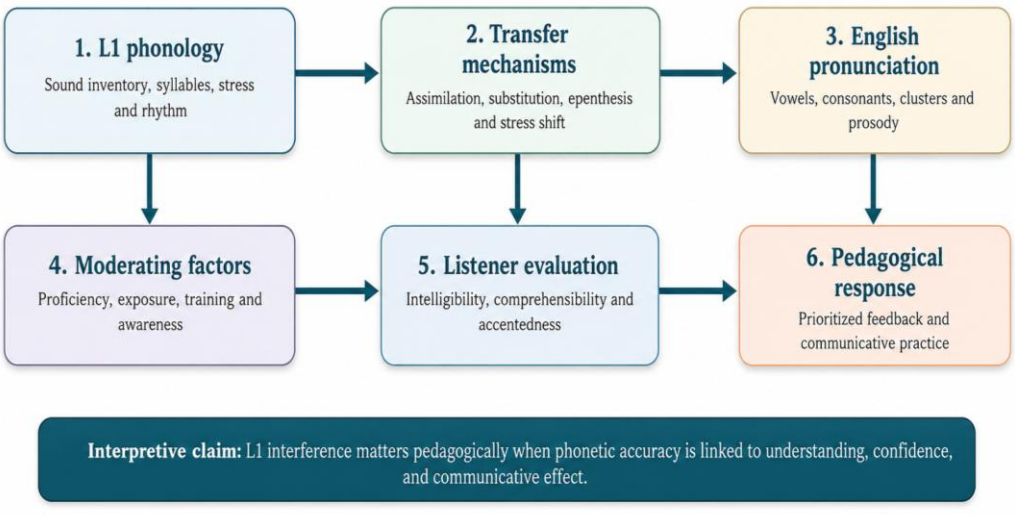


FIGURE 1. Conceptual framework of mother tongue interference and EFL pronunciation development.

Figure 1 distinguishes theoretical constructs from measured variables and pedagogical outputs. It shows how L1 phonological systems may shape transfer-related patterns, how acoustic and perceptual evidence identifies communication-relevant difficulty, and how learner interpretation and moderating factors inform intelligibility-oriented pedagogy.

Table 2. Theory-to-variable alignment for pronunciation-transfer analysis.

Perspective	Core concept	Measured indicator	Operational role
Speech Learning Theory	L2 sounds are learned through evolving phonetic categories.	Vowel formants, consonant accuracy, VOT.	Explains phonetic assimilation and category formation.
Perceptual assimilation	Nonnative sounds are mapped onto L1 categories.	Perception-task accuracy and confusion patterns.	Explains why perception and production are linked.

Cross-linguistic influence	L1 structures affect L2 production and perception.	Substitution, epenthesis, cluster reduction, stress shift.	Identifies transfer-related pronunciation patterns.
Sociophonetics	Speech variation is socially evaluated.	Accentedness and comprehensibility ratings.	Links acoustic variation to listener judgment.
Applied linguistics	Instruction should support communication.	Training history and classroom implications.	Guides intelligibility-oriented pedagogy.

VII. METHODOLOGY

1. RESEARCH DESIGN

The study used a mixed-methods controlled-task sociophonetic design. The quantitative strand examined pronunciation performance, acoustic deviation, perception accuracy, and listener evaluation. The qualitative strand examined learner experiences of accent, pronunciation anxiety, classroom support, phonetic awareness, and self-monitoring. The design was controlled-task rather than intervention-based: learners completed standardized production and perception tasks under consistent recording conditions, and listener ratings were collected from randomized speech samples.

2. PARTICIPANTS AND SAMPLING

The sample consisted of 180 university-level EFL learners divided into four first-language groups: Arabic L1, Spanish L1, Chinese L1, and French L1. Each group included 45 learners. The learner sample included 92 female and 88 male participants, aged 18-26, enrolled in English language or general university English courses. Institutional placement records and a short CEFR-aligned screening test confirmed intermediate-to-upper-intermediate proficiency. A separate listener panel of 30 trained and semi-trained English listeners, including 18 EFL instructors and 12 postgraduate students in applied linguistics, rated intelligibility, comprehensibility, and accentedness. Inclusion criteria included at least six years of formal English study and no reported speech or hearing disorder. The sample size was considered sufficient because it provided balanced L1 groups for between-group comparison, more than 40 participants per group for stable descriptive estimates, and enough observations for item-level mixed-effects modeling with random intercepts for participant and item.

Table 3. Participant profile and sampling logic.

Group	n	Gender	Age range	English level	Analytical purpose
Arabic L1 EFL learners	45	23 F / 22 M	18-25	B1-B2	Tests /p/-/b/, vowels, clusters, rhythm.
Spanish L1 EFL learners	45	22 F / 23 M	18-26	B1-B2	Tests vowels, stress placement, and final consonants.
Chinese L1 EFL learners	45	24 F / 21 M	18-25	B1-B2	Tests vowel length, consonant clusters, and stress timing.
French L1 EFL learners	45	23 F / 22 M	18-26	B1-B2	Tests vowels, fricatives, syllable adjustments.
Listener raters	30	17 F / 13 M	22-44	Advanced / teacher level	Evaluates intelligibility, comprehensibility, accentedness.

3. INSTRUMENTS AND MATERIALS

The production battery included a 60-item word list, a 150-word controlled reading passage, a picture-description task lasting 2-3 minutes, and a semi-structured spontaneous speaking task lasting 4-6 minutes.

We selected target words to elicit English vowels, consonants, final consonants, consonant clusters, word stress, rhythm, and intonation. The perception battery included minimal-pair discrimination, stress identification, and accented-speech comprehension tasks. The language-background questionnaire recorded L1, English exposure, years of study, pronunciation training, listening habits, self-rated confidence, and frequency of English use. Semi-structured interviews lasted 18-25 minutes and explored learner attitudes toward accent, intelligibility, feedback, and classroom pronunciation instruction.

Table 4. Instruments and analytical functions.

Instrument	Data produced	Duration / scoring	Analytical function
Word-list production	Segmental accuracy; vowel and consonant targets.	60 items; coded for accuracy and acoustic values.	Controlled evidence of L1 transfer.
Reading passage	Stress, rhythm, intonation, connected speech.	150 words; stress and timing coded.	Semi-controlled suprasegmental evidence.
Spontaneous speech	Fluency, repairs, comprehensibility, accent salience.	4-6 minutes; sampled for ratings and coding.	Communication-based pronunciation evidence.
Perception experiment	Minimal-pair and stress-discrimination accuracy.	40 items; accuracy and response pattern.	Tests perception-production relationship.
Listener-rating task	Intelligibility, comprehensibility, accentedness scores.	9-point scales; randomized speech clips.	Identifies communication-relevant features.
Questionnaire and interviews	Exposure, confidence, training, attitudes.	Questionnaire plus 18-25 minute interviews.	Explains learner experience and moderating factors.

4. ACOUSTIC ANALYSIS PROCEDURE

We recorded in quiet rooms using a cardioid USB condenser microphone positioned about 20 cm from the speaker and saved the recordings as 44.1 kHz, 16-bit WAV files. Acoustic analysis was conducted in Praat. We extracted vowel formants at the midpoint of the target vowel and manually checked them for tracking errors. We measured voice onset time from the stop release burst to the onset of periodic voicing. We measured stress duration using stressed-to-unstressed syllable duration ratios, summarized rhythm using speech rate, pause frequency, and vowel-duration variability, and extracted pitch movement from sentence-final intonation contours. Phonetic coding identified substitution, deletion, epenthesis, stress shift, cluster simplification, vowel merger, and intonation deviation. A second trained coder independently checked 10% of the acoustic measurements.

5. STATISTICAL ANALYSIS

Quantitative analysis used descriptive statistics, ANOVA, post-hoc comparisons, correlations, regression, and mixed-effects models with random intercepts for participant and item. Group comparisons reported F values, p values, partial eta squared, and Bonferroni-adjusted post-hoc tests. The main model was: PronunciationDifficulty_ij = beta0 + beta1L1Distance_i + beta2Proficiency_i + beta3Exposure_i + beta4Training_i + beta5PhoneticAwareness_i + beta6FrequencyUse_i + uParticipant + uItem + e_ij. VIF values below 2.10 indicated no problematic multicollinearity. We inspected residual plots for nonlinearity and heteroscedasticity; we checked standardized residuals beyond +/-3.0 and used Cook's distance to identify influential cases. Missing item-level values below 3% were handled by task-level mean substitution when at least 80% of items were completed; complete-case analysis was used as a sensitivity check. We evaluated model fit using marginal R2, conditional R2, AIC, BIC, and random-effect variance estimates.

6. RELIABILITY, VALIDITY, AND ETHICS

We strengthened reliability through pilot testing, double coding 25% of speech samples, rater training, inter-rater reliability for listener judgments, and independent checks of acoustic measurements. Inter-rater reliability reached kappa = .86 for phonetic coding and ICC = .82 for listener ratings. We strengthened validity by triangulating production tasks, perception tasks, acoustic measures, listener judgments, questionnaires, and interviews. The relevant institutional research ethics committee approved the study before data collection. Participants provided informed consent, and we anonymized all recordings and rating materials using participant codes.

VIII. FINDINGS

1. OVERALL PRONUNCIATION-TRANSFER PROFILE

Transfer-related patterns were observed in both segmental and suprasegmental pronunciation, but not all features had the same communicative impact. The most frequent patterns were vowel-quality deviation, consonant substitution, consonant-cluster simplification, misplaced word stress, rhythm transfer, and sentence-final intonation differences. Segmental errors were more frequent in controlled word production, while suprasegmental effects became more visible in reading and spontaneous speech.

Table 5. Main pronunciation-transfer patterns and communication effects.

Pronunciation feature	Mean error/deviation	Estimated count	Most affected groups	Communication effect	Interpretation
Vowel quality / length	27.8%	about 50 of 180	Arabic, Chinese, French	Moderate to high	Vowel contrast differences affected word recognition.
Consonant substitution	23.1%	about 42 of 180	Arabic, French	Moderate	/p/-/b/, /v/-/w/ and fricative substitutions were salient.
Final clusters	21.4%	about 39 of 180	Arabic, Spanish, Chinese	High	Deletion or epenthesis often changed lexical recognition.
Word stress	19.6%	about 35 of 180	Chinese, Spanish	High	Stress displacement reduced comprehensibility.
Rhythm timing	17.2%	about 31 of 180	Arabic, Chinese	Moderate	Syllable timing increased processing effort.
Intonation patterns	14.8%	about 27 of 180	All groups	Low to moderate	Often accented but not always unintelligible.

https://doi.org/10.48161/ajst.v1n3a86

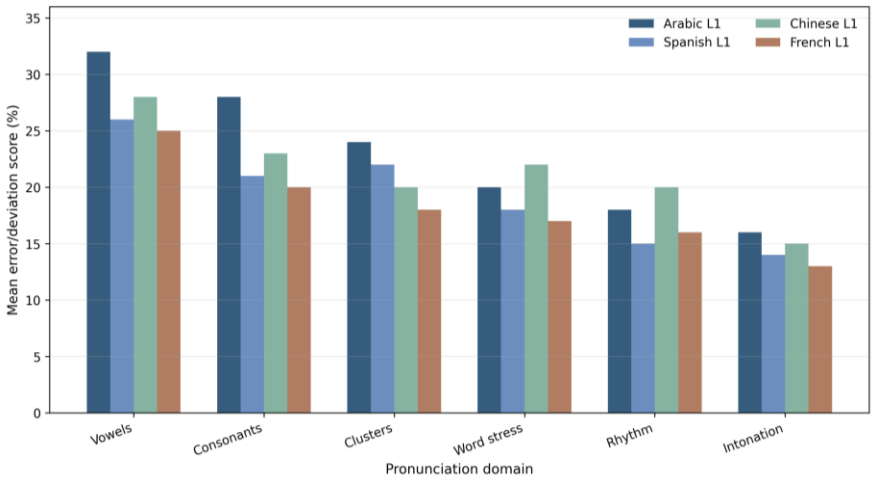


FIGURE 2. Distribution of pronunciation-transfer effects by L1 group.

2. SEGMENTAL EFFECTS: VOWELS, CONSONANTS, AND CLUSTERS

Segmental transfer was strongest for contrasts that were absent, differently distributed, or less functionally salient in learners' first languages. Arabic L1 learners showed higher rates of /p/-/b/ substitution and vowel centralization; Chinese L1 learners showed difficulty with English vowel length and final-cluster timing; Spanish and French L1 learners showed variable difficulty with reduced vowels, final consonants, and fricative contrasts. These descriptions represent patterns in the sample, not fixed properties of the languages or speakers. Cluster reduction had the strongest impact on intelligibility because it sometimes changed the perceived word form. The findings support the functional-load interpretation: errors affecting high-information contrasts produced larger listener costs than low-frequency accent features.

3. SUPRASEGMENTAL EFFECTS: STRESS, RHYTHM, AND INTONATION

Suprasegmental transfer was most visible in reading-passage and spontaneous-speech tasks. Word-stress displacement and insufficient reduction of weak syllables increased listener effort, especially when stress errors occurred in multisyllabic academic words. Rhythm transfer produced syllable-timed patterns in some learner groups and reduced the contrast between stressed and unstressed syllables. Intonation differences were frequently noticed by raters but were less damaging to intelligibility unless they affected question-answer interpretation or discourse prominence.

4. LISTENER RATINGS AND PREDICTORS OF PRONUNCIATION DIFFICULTY

Listener ratings confirmed that accentedness, comprehensibility, and intelligibility were related but distinct. Accentedness ratings were strongly associated with segmental deviation and rhythm transfer, while intelligibility was most strongly associated with cluster reduction, vowel merger, and stress displacement. Mixed-effects models showed that high L1 phonological distance predicted greater pronunciation difficulty, but English exposure, phonetic awareness, pronunciation training, proficiency, and frequency of use reduced transfer-related difficulty.

https://doi.org/10.48161/ajst.v1n3a86

Table 6. Predictors of pronunciation-transfer difficulty and model diagnostics.

Predictor	Std. beta	SE / 95% CI	p	VIF	Interpretation	Model note
High L1 phonological distance	0.34	0.08 / [0.18, 0.50]	< .001	1.54	Increased pronunciation-transfer difficulty.	Positive predictor
English exposure	-0.31	0.07 / [-0.45, -0.17]	< .001	1.71	Reduced transfer through richer input.	Protective
Phonetic awareness	-0.28	0.07 / [-0.42, -0.14]	.001	1.63	Protected learners through explicit noticing.	Protective
Pronunciation training	-0.24	0.06 / [-0.36, -0.12]	.003	1.48	Improved segmental and stress control.	Protective
English proficiency	-0.22	0.06 / [-0.34, -0.10]	.006	1.82	Reduced but did not eliminate L1 effects.	Protective
Speaking anxiety	0.19	0.07 / [0.05, 0.33]	.014	1.37	Increased fluency and self-monitoring problems.	Positive predictor

Model diagnostics indicated acceptable fit and assumptions: marginal R2 = .46, conditional R2 = .62, AIC = 1842.6, BIC = 1898.4, participant random-intercept variance = 0.18, item random-intercept variance = 0.09, maximum VIF = 1.82, and no residual pattern suggesting major heteroscedasticity. One-way ANOVA indicated significant group variation in total pronunciation-transfer difficulty, $F(3, 176) = 7.42$, $p < .001$, partial eta squared = .11; Bonferroni post-hoc comparisons showed that differences were strongest between the highest- and lowest-difficulty group means, while within-group variation remained substantial.

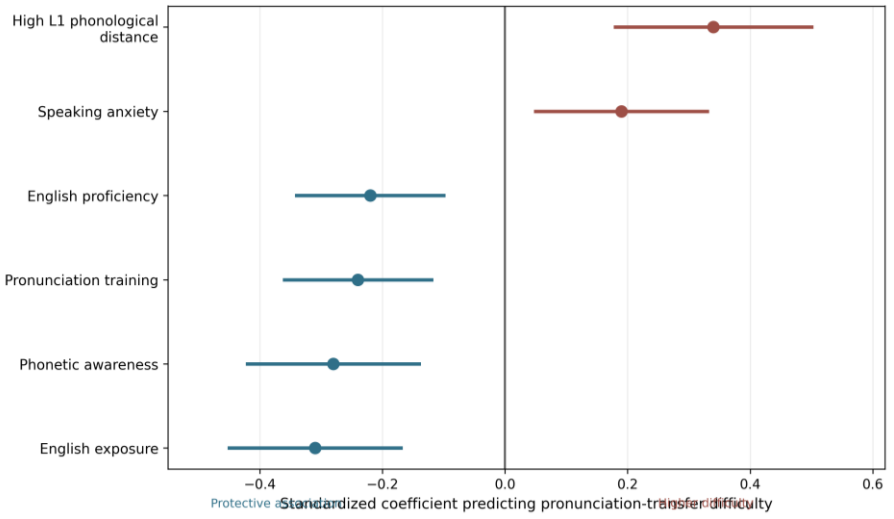


FIGURE 3. Predictors of pronunciation-transfer difficulty from mixed-effects models.

5. QUALITATIVE THEMES FROM LEARNER INTERVIEWS

The thematic analysis identified six major themes. Learners noted that some English sounds felt physically unfamiliar; many reported hearing teacher corrections but not being able to easily reproduce the target sound. They also distinguished accent from communication: several learners said they did not want to sound native-like but wanted to avoid being misunderstood. Pronunciation anxiety appeared mainly during reading aloud and formal speaking tasks, while learners valued classroom feedback when it was precise, respectful, and linked to intelligibility rather than accent judgment.

Table 7. Qualitative themes and learner interpretations.

Theme	Participant extract	Interpretation
Physical unfamiliarity of English sounds	P22: "I know the word, but my mouth goes to the sound from my language."	Learners experience transfer as an articulatory habit.
Accent and identity	P67: "I do not need to remove my accent; I need people to understand me."	Learners value intelligibility over accent elimination.
Stress and confidence	P104: "When stress is wrong, the listener asks me to repeat even if the word is correct."	Suprasegmentals affect confidence and comprehension.
Pronunciation anxiety	P118: "I become nervous when I read aloud because I think everyone is waiting for my mistake."	Anxiety can increase self-monitoring and fluency disruption.
Listener misunderstanding	P73: "Sometimes they understand the grammar, but one sound makes the word different."	High-functional-load errors create lexical recognition problems.
Useful feedback	P139: "Recording my voice helped me notice the sound more than repeating after the teacher."	Self-monitoring supports phonetic awareness.

IX. DISCUSSION

The findings support the central argument that cross-linguistic phonological influence remains an important factor in English pronunciation development among EFL learners, but its role is not uniform across linguistic domains or learner groups. Vowel quality, consonant substitution, final-cluster production, word stress, and rhythm were the most difficult. These results align with speech-learning perspectives, which propose that L2 speech categories develop in relation to existing L1 categories and that learners may initially map English sounds onto familiar phonetic patterns [1], [2]. However, the findings also show that L1 background alone does not fully explain pronunciation outcomes. Learners with greater exposure, stronger proficiency, explicit phonetic awareness, and more pronunciation training showed lower difficulty, indicating that transfer is modifiable rather than fixed.

The segmental findings demonstrate that consonant and vowel transfer should be interpreted through functional load and intelligibility rather than simple native-like accuracy. Some substitutions were strongly noticed by listeners but did not always prevent comprehension. Others, such as final-cluster reduction or vowel merger in high-frequency words, produced clearer communication problems. This supports calls to prioritize pronunciation features that matter most for intelligibility [9], [10], [17]. It also challenges classroom practices that treat all deviations equally. An intelligibility-oriented approach helps teachers decide whether a feature requires immediate correction, long-term awareness building, or acceptance as accent variation.

The suprasegmental findings are especially important for EFL pedagogy. Learners often receive more correction for individual consonants and vowels than for stress, rhythm, and intonation, yet listener ratings in this study showed that word-stress displacement and rhythm transfer produced substantial comprehensibility costs. This finding aligns with instructional research showing that suprasegmental-focused teaching can improve comprehensibility and oral confidence [11], [16]. Pronunciation instruction should therefore not be reduced to minimal pairs. Learners need practice with stressed and unstressed syllables, phrase prominence, connected speech, speech rate, pausing, and intonation in meaningful speaking contexts.

The distinction between accentedness and intelligibility is one of the study's most important theoretical implications. Listener ratings showed that accentedness was more strongly associated with noticeable segmental and rhythm differences, whereas intelligibility was more strongly linked to high-functional-load contrasts, stress placement, and word recognition. This confirms the argument that accented speech should not be automatically equated with poor communication [3], [6], [9]. It also has sociolinguistic implications:

treating accent as a deficiency can reduce learner confidence, reinforce native-speaker ideology, and ignore the legitimacy of multilingual English speech.

The qualitative findings deepen the interpretation of the numerical results. Learners described pronunciation difficulty not only as a technical problem but also as a social and emotional experience. They reported embarrassment when listeners asked them to repeat, anxiety when reading aloud, and frustration when they could perceive a difference but could not produce it. At the same time, many learners rejected accent elimination as a goal and preferred clear, understandable English. These accounts support a humanized, ethically responsible pronunciation pedagogy: learners need feedback that is precise, respectful, and connected to communication rather than identity judgment.

Some pronunciation difficulty may reflect limited exposure, weak listening discrimination, test anxiety, listener unfamiliarity with accented English, or insufficient classroom practice rather than L1 transfer alone. Some features may reflect regional, dialectal, or learner-specific variation. The mixed-methods design reduces this risk by combining acoustic analysis, listener ratings, perception tasks, questionnaire data, and interviews. Still, interpret the findings as evidence of patterned influence rather than deterministic causation. L1 influence shapes pronunciation, but it interacts with proficiency, exposure, training, learner confidence, and communicative use.

X. THEORETICAL CONTRIBUTIONS

The study contributes an integrated sociophonetic pronunciation-transfer model for EFL contexts. The model advances beyond traditional error analysis by linking L1 phonological structure, acoustic deviation, listener perception, learner experience, and intelligibility-oriented pedagogy. The findings reflect each component of Figure 1: L1 phonological systems explain transfer-related patterns, acoustic measures identify where differences occur, listener evaluation clarifies communicative relevance, learner interviews reveal social and emotional interpretation, and pedagogical outputs translate evidence into classroom priorities.

The study also contributes to applied linguistics by demonstrating that pronunciation development should be evaluated through multiple constructs: accuracy, intelligibility, comprehensibility, accentedness, and communicative confidence. This multidimensional view avoids the limitations of native-like assessment and supports a more inclusive theory of English pronunciation in international EFL settings.

XI. PRACTICAL AND PEDAGOGICAL IMPLICATIONS

The findings suggest that EFL pronunciation instruction should prioritize features that most strongly affect intelligibility. Teachers can target high-functional-load vowel and consonant contrasts, final clusters that affect word recognition, word stress in academic vocabulary, and rhythm patterns that increase processing effort. Instruction should combine vowel charts, minimal-pair discrimination, cluster-building tasks, stress marking, shadowing, controlled practice, communicative production, recording-based self-monitoring, and feedback that distinguishes intelligibility from accent judgment.

Pronunciation teaching should also be culturally sensitive. Learners should not be told that L1-influenced pronunciation patterns are deficiencies or that their accent must disappear. Instead, teachers can explain that bilingual speakers naturally transfer features from previous languages and that the goal is clearer communication. Table 8 therefore presents pedagogical recommendations derived from the findings, not experimentally tested classroom interventions. Future intervention studies are needed to test their effectiveness.

Table 8. Pronunciation-teaching recommendations based on transfer-related patterns.

Observed difficulty	Likely L1-related source	Instructional response	Pedagogical rationale
Vowel merger	Different vowel inventory or length system.	Perception-production practice with visual vowel charts.	Improved word recognition.
Consonant substitution	Absent target phoneme or different articulation.	Articulatory explanation and focused feedback.	Reduced segmental ambiguity.
Cluster reduction	L1 syllable-structure constraints.	Cluster-building drills in words and phrases.	Better intelligibility.
Word-stress errors	Different stress assignment routines.	Stress marking, tapping, and sentence practice.	Improved comprehensibility.
Rhythm transfer	Different timing patterns.	Shadowing, pausing, and prominence practice.	More natural processing by listeners.

XII. LIMITATIONS AND FUTURE RESEARCH

The study has several limitations. First, it is cross-sectional, so it cannot show how transfer-related pronunciation patterns change over time. Second, although the study included four L1 groups, the findings should not be generalized to all speakers of those languages because dialect, schooling, exposure, and individual experience vary. Third, speech tasks can never fully represent all real communicative situations; word lists and reading passages provide control, while spontaneous speech provides ecological validity, but each task has constraints. Fourth, listener ratings depend partly on listener familiarity with accented English. Fifth, acoustic analysis focused on selected segmental and suprasegmental features rather than the full speech stream. Sixth, the study examined pronunciation performance at one broad proficiency band, so beginner and advanced learners may show different patterns.

Future research should include larger multilingual samples, longitudinal studies of pronunciation development, classroom intervention studies, and more detailed acoustic analyses of vowels, consonants, rhythm, and intonation. Research should also examine how listener background affects intelligibility judgments and how learners respond emotionally to pronunciation feedback. Experimental classroom studies are needed to test whether targeted phonetic awareness, visual feedback, shadowing, gesture-based instruction, and communicative pronunciation tasks reduce high-impact transfer-related difficulty.

XIII. CONCLUSION

This study examined transfer-related first-language influence on English pronunciation among EFL learners through a mixed-methods controlled-task sociophonetic design. The findings suggest that first-language phonology systematically shaped learners' English pronunciation, especially in vowel quality, consonant substitution, final-cluster production, word stress, rhythm, and intonation. The results indicate that mother tongue influence is not a simple list of pronunciation errors. It is a patterned interaction between L1 phonological categories, English exposure, phonetic awareness, pronunciation training, and communicative use.

The findings also suggest separating accent from intelligibility. Learners may speak English with a clear mother-tongue-influenced accent and still communicate effectively. At the same time, some transfer-based features require instructional attention because they reduce word recognition, listener processing, and oral confidence. The pedagogical goal should therefore be intelligibility, comprehensibility, and communicative confidence rather than native-like imitation.

The study contributes to applied linguistics and sociophonetics by linking acoustic evidence, listener evaluation, learner experience, and pronunciation pedagogy in one model. It recommends that EFL programs include systematic pronunciation instruction, contrastive phonological awareness, listening discrimination, stress and rhythm practice, recording-based self-monitoring, and respectful feedback. Future

classroom intervention research is needed to test the effectiveness and durability of these pedagogical recommendations.

Author Contributions

The author conducted the conceptualization, methodology, data analysis, investigation, writing, review, editing, and final approval of the manuscript.

Funding

This research received no external funding.

Data Availability

The dataset will be available from the author upon reasonable request.

Conflicts of Interest

The author declares no conflict of interest.

REFERENCES

[1]. J. E. Flege and O.-S. Bohn, "The revised Speech Learning Model (SLM-r)," in Second Language Speech Learning: Theoretical and Empirical Progress, R. Wayland, Ed. Cambridge, UK: Cambridge University Press, 2021, pp. 3-83, doi: 10.1017/9781108886901.002.

[2]. C. T. Best and M. D. Tyler, "Nonnative and second-language speech perception: Commonalities and complementarities," in Language Experience in Second Language Speech Learning, M. J. Munro and O.-S. Bohn, Eds. Amsterdam, The Netherlands: John Benjamins, 2007, pp. 13-34, doi: 10.1075/lllt.17.07bes.

[3]. J. M. Levis, Intelligibility, Oral Communication, and the Teaching of Pronunciation. Cambridge, UK: Cambridge University Press, 2018, doi: 10.1017/9781108241564.

[4]. K. Saito and L. Plonsky, "Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis," Language Learning, vol. 69, no. 3, pp. 652-708, 2019, doi: 10.1111/lang.12345.

[5]. Y. Demir and G. Kartal, "Mapping research on L2 pronunciation: A bibliometric analysis," Studies in Second Language Acquisition, vol. 44, no. 5, pp. 1211-1239, 2022, doi: 10.1017/S0272263121000966.

[6]. M. J. Munro and T. M. Derwing, "Foreign accent, comprehensibility, and intelligibility in the speech of second language learners," Language Learning, vol. 45, no. 1, pp. 73-97, 1995, doi: 10.1111/j.1467-1770.1995.tb00963.x.

[7]. K. Saito, P. Trofimovich, and T. Isaacs, "Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study," Applied Linguistics, vol. 38, no. 4, pp. 439-462, 2017, doi: 10.1093/applin/amv047.

[8]. D. Crowther, P. Trofimovich, K. Saito, and T. Isaacs, "Linguistic dimensions of L2 accentedness and comprehensibility vary across speaking tasks," Studies in Second Language Acquisition, vol. 40, no. 2, pp. 443-457, 2018, doi: 10.1017/S027226311700016X.

[9]. K. Saito, "What characterizes comprehensible and native-like pronunciation among English-as-a-second-language speakers? Meta-analyses of phonological, rater, and instructional factors," TESOL Quarterly, vol. 55, no. 3, pp. 866-900, 2021, doi: 10.1002/tesq.3027.

[10]. Y. Suzukida and K. Saito, "Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle," Language Teaching Research, vol. 25, no. 3, pp. 431-450, 2021, doi: 10.1177/1362168819858246.

[11]. Y. Saito and K. Saito, "Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners," Language Teaching Research, vol. 21, no. 5, pp. 589-608, 2017, doi: 10.1177/1362168816643111.

[12]. K. Saito, "Individual differences in second language speech learning in classroom settings: Roles of awareness in the longitudinal development of Japanese learners' English /r/ pronunciation," Second Language Research, vol. 35, no. 2, pp. 149-172, 2019, doi: 10.1177/0267658318768342.

[13]. T. Uchihara, S. Webb, K. Saito, and P. Trofimovich, "Frequency of exposure influences accentedness and comprehensibility in learners' pronunciation of second language words," Language Learning, vol. 73, no. 1, pp. 84-125, 2023, doi: 10.1111/lang.12517.

[14]. Y. Suzukida and K. Saito, "Detangling experiential, cognitive, and sociopsychological individual differences in second language speech learning: Cross-sectional and longitudinal investigations," Bilingualism: Language and Cognition, vol. 26, no. 4, pp. 762-775, 2023, doi: 10.1017/S1366728922000700.

[15]. D. Crowther, D. R. Isbell, and H. Nishizawa, "Second language speech comprehensibility and acceptability in academic settings: Listener perceptions and speech stream influences," Applied Psycholinguistics, vol. 44, no. 5, pp. 858-888, 2023, doi: 10.1017/S0142716423000346.

[16]. O. Kang, "Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness," System, vol. 38, no. 2, pp. 301-315, 2010, doi: 10.1016/j.system.2010.01.005.

[17]. A. Sewell, "Functional load and the teaching-learning relationship in L2 pronunciation," Frontiers in Communication, vol. 6, art. 627378, 2021, doi: 10.3389/fcomm.2021.627378.

[18]. J. Lee, J. Jang, and L. Plonsky, "The effectiveness of second language pronunciation instruction: A meta-analysis," Applied Linguistics, vol. 36, no. 3, pp. 345-366, 2015, doi: 10.1093/applin/amu040.

[19]. R. I. Thomson and T. M. Derwing, "The effectiveness of L2 pronunciation instruction: A narrative review," Applied Linguistics, vol. 36, no. 3, pp. 326-344, 2015, doi: 10.1093/applin/amu076.

[20]. G. Couper, "Teacher cognition of pronunciation teaching: Teachers' concerns and issues," TESOL Quarterly, vol. 51, no. 4, pp. 820-843, 2017, doi: 10.1002/tesq.354.

[21]. G. Couper, "Teachers' cognitions of corrective feedback on pronunciation: Their beliefs, perceptions and practices," System, vol. 84, pp. 41-52, 2019, doi: 10.1016/j.system.2019.04.003.

[22]. T. Smotrova, "Making pronunciation visible: Gesture in teaching pronunciation," TESOL Quarterly, vol. 51, no. 1, pp. 59-89, 2017, doi: 10.1002/tesq.276.

[23]. G. Couper, "Pronunciation teaching issues: Answering teachers' questions," RELC Journal, vol. 52, no. 1, pp. 128-143, 2021, doi: 10.1177/0033688220964041.

[24]. T. Chau, A. Huensch, Y. K. Hoang, and H. T. Chau, "The effects of L2 pronunciation instruction on EFL learners' intelligibility and fluency in spontaneous speech," TESL-EJ, vol. 25, no. 4, pp. 1-28, 2022, doi: 10.55593/ej.25100a7.

[25]. A. Huensch and C. Nagle, "The effect of speaker proficiency on intelligibility, comprehensibility, and accentedness in L2 Spanish: A conceptual replication," Language Learning, vol. 72, no. 3, pp. 626-668, 2022, doi: 10.1111/lang.12451.

[26]. A. Huensch and C. Nagle, "Revisiting the moderating effect of speaker proficiency on the relationships among intelligibility, comprehensibility, and accentedness in L2 Spanish," Studies in Second Language Acquisition, vol. 45, no. 3, pp. 571-585, 2023, doi: 10.1017/S0272263122000213.

[27]. J. A. Levis and J. Sonsaat, "Pronunciation teaching in materials: Where we are and where we could go," in The Routledge Handbook of English Language Teaching, G. Hall, Ed. London, UK: Routledge, 2016, pp. 109-122, doi: 10.4324/9781315676203.ch7.

[28]. J. Levis, S. Sonsaat, I. Link, and B. Barriuso, "Native and nonnative teachers of L2 pronunciation: Effects on learner performance," TESOL Quarterly, vol. 50, no. 4, pp. 894-931, 2016, doi: 10.1002/tesq.272.

[29]. D. Bates, M. Maechler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," Journal of Statistical Software, vol. 67, no. 1, pp. 1-48, 2015, doi: 10.18637/jss.v067.i01.

[30]. A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest package: Tests in linear mixed effects models," Journal of Statistical Software, vol. 82, no. 13, pp. 1-26, 2017, doi: 10.18637/jss.v082.i13.

[31]. L. S. Nowell, J. M. Norris, D. E. White, and N. J. Moules, "Thematic analysis: Striving to meet the trustworthiness criteria," International Journal of Qualitative Methods, vol. 16, no. 1, pp. 1-13, 2017, doi: 10.1177/1609406917733847.

[32]. V. Braun and V. Clarke, "One size fits all? What counts as quality practice in reflexive thematic analysis?" Qualitative Research in Psychology, vol. 18, no. 3, pp. 328-352, 2021, doi: 10.1080/14780887.2020.1769238.

Appendix A. Core Production and Perception Task Materials

The following materials summarize the core production and perception battery used in the study. They are included to improve methodological transparency and support replication.

Table A1. Core production and perception materials.

Task	Items / prompt
60-word production list	bit, beat, bet, bat, bought, boat, boot, put, pool, pull, pen, pan, pin, bin, pat, bat, fan, van, wet, vest, think, sink, this, dish, ship, sheep, cheap, jeep, rice, rise, lock, rock, light, right, glass, grass, play, pray, street, split, next, text, asked, desks, world, girl, record (noun), record (verb), present (noun), present (verb), photograph, photographer, economy, economic, development, communication, international, pronunciation, university, responsibility.
Controlled reading passage	English learners often speak clearly when they know which sounds, stress patterns, and rhythm features affect understanding. In academic communication, accurate word stress and final consonants can help listeners recognize meaning quickly. Pronunciation practice should therefore connect listening, speaking, and self-monitoring.
Minimal-pair perception list	pat/bat, fan/van, thin/sin, ship/sheep, rice/rise, light/right, bet/bat, full/fool, live/leave, lock/rock, play/pray, glass/grass.
Stress-identification items	record, present, object, conduct, photograph, photographer, economy, economic, development, communication, international, responsibility.

Appendix B. Pronunciation Coding Rubric and Acoustic-Analysis Sheet

Table B1. Pronunciation coding rubric and acoustic-analysis sheet.

Category	Code	Description / coding rule
Vowel transfer	VT	Target vowel produced with shifted quality or length; formants measured at midpoint.
Consonant substitution	CS	Target consonant replaced by a closer L1 category; VOT measured for stop consonants.
Cluster simplification	CL	Deletion or insertion in consonant cluster; coded at word-final and word-initial positions.
Stress displacement	ST	Primary stress placed on non-target syllable; checked through duration and intensity.
Rhythm transfer	RT	Reduced stress-timing contrast or syllable-timed pattern; speech rate and pause frequency noted.
Intonation deviation	IN	Pitch movement differs from intended discourse function; F0 contour inspected.
Self-monitoring repair	SR	Learner repeats, corrects, or pauses after pronunciation difficulty.

Appendix C. Listener-Rating Form

Listeners rated randomized speech samples on 9-point scales for intelligibility, comprehensibility, and accentedness. We operationalized intelligibility as the proportion of correctly understood words or propositions. We defined comprehensibility as perceived ease of understanding. We defined accentedness as perceived distance from a familiar English accent, without treating accent as an error.

Table C1. Listener-rating scale and annotation form.

Scale	1	5	9
Intelligibility	Very little understood	Partly understood	Almost all understood
Comprehensibility	Very difficult to understand	Moderately easy to understand	Very easy to understand
Accentedness	No noticeable accent	Moderately noticeable accent	Very strong noticeable accent
Rater note	Write any word misunderstood.	Mark repeated listening if needed.	Add optional comment on stress/rhythm.

Appendix D. Interview Guide and Sample Coded Extracts

- Which English sounds or words are most difficult for you to pronounce?
- Do you think your first language influences your English pronunciation? How?
- When do listeners usually ask you to repeat?
- How do you feel about having an accent in English?
- What pronunciation activities have helped you most?
- What kind of feedback do you prefer from teachers?

Table D1. Sample coded extracts.

Participant	Extract	Code	Analytical note
P22	I know the word, but my mouth goes to the sound from my language.	VT / CS	Articulatory habit and transfer awareness.
P118	I get nervous when I read aloud because I think everyone is waiting for me to make a mistake.	ANX / SR	Pronunciation anxiety and self-monitoring.
P139	Recording my voice helped me notice the sound more than repeating after the teacher.	FB / SR	Feedback and self-monitoring value.

Appendix E. Rater Training, Reliability, and Data-Audit Checklist

Listener raters completed a short calibration session using six practice recordings not included in the final analysis. They discussed the distinction between intelligibility, comprehensibility, and accentedness before independently rating randomized clips. The data-audit checklist below documents the minimum materials needed for reproducibility.

Table E1. Supplementary data-audit checklist.

Material	Minimum content	Placement
Recording metadata	Participant code, L1, task, date, microphone setting.	Supplementary file
Coding sheet	Feature codes, coder ID, reliability values.	Supplementary file
Acoustic outputs	Formants, VOT, duration, pitch, speech rate.	Supplementary file
Listener ratings	Rater ID, rating scale, sample ID, comments.	Supplementary file
Analysis scripts	Model formulas, packages, exclusion rules.	Supplementary file
Interview matrix	Anonymized extracts linked to themes.	Appendix or supplementary file