

## **Integrating Pronunciation into Communicative Language Teaching: A Systematic Review of Pedagogical Evidence**

**Awad Alshehri\***

Imam Mohammad Ibn Saud Islamic University (IMSIU),  
College of Languages and Translation.

### **Abstract**

This systematic review examines how integrating pronunciation instruction into Communicative Language Teaching (CLT) influences learning outcomes in English as a Second/Foreign Language (ESL/EFL) contexts. Following PRISMA guidelines, we searched ERIC, LLBA, PsycINFO, Scopus, and Web of Science for studies published between 1980 and 2024, screening and appraising eligible studies for methodological quality. Thirty-seven studies met the inclusion criteria. Findings indicate that CLT-integrated pronunciation instruction consistently improves intelligibility and comprehensibility more than decontextualized or form-focused approaches, especially when paired with explicit feedback, task-based activities, and at least 10 hours of instruction. Gains were stronger for suprasegmental features (e.g., stress, rhythm, intonation) than for segmentals alone, though evidence on long-term retention and learner variables (e.g., L1, proficiency) remains limited. Overall, embedding explicit pronunciation teaching within communicative tasks yields more robust outcomes than isolated drills, suggesting the need for standardized measures, longitudinal designs, and pre-registered protocols in future research. Language programs should prioritize intelligibility-focused targets, integrate feedback into communicative tasks, and align assessments with communicative goals.

**Keywords:** Communicative Language Teaching (CLT); Comprehensibility; Corrective Feedback; ESL/EFL Pedagogy; Intelligibility; Pronunciation Instruction; Second Language Acquisition (SLA); Suprasegmentals; Systematic Review; Task-based Learning

### **1. Introduction**

Pronunciation has long been recognized as a critical yet often underemphasized component of second language instruction. In English as a Second or Foreign Language (ESL/EFL) contexts, learners frequently report difficulties with segmental and suprasegmental features that affect intelligibility and comprehensibility. While Communicative Language Teaching (CLT) has become the dominant paradigm in language pedagogy, the role of explicit pronunciation instruction within this framework remains contested. Traditional approaches often separated pronunciation from communicative practice, focusing on decontextualized drills or native-like accuracy, whereas more recent perspectives emphasize intelligibility and communicative competence as primary goals.

Over the past four decades, research has explored a range of instructional models integrating pronunciation into CLT, from task-based learning and focus-on-form activities to technology-enhanced feedback and strategy training. Several studies suggest that embedding pronunciation instruction into communicative tasks can yield greater gains in speech clarity and learner confidence than isolated practice alone. However, findings remain scattered across contexts,

target features, and instructional designs, with no consensus on the most effective approaches or conditions.

Despite multiple narrative reviews, there has been no systematic synthesis of empirical evidence examining how CLT-integrated pronunciation instruction affects learner outcomes. Such a synthesis is essential for informing evidence-based pedagogy, guiding curriculum design, and shaping teacher training practices. This review therefore aims to (a) map the empirical research on pronunciation within CLT frameworks, (b) evaluate the pedagogical effectiveness of different instructional approaches, and (c) identify gaps and directions for future research.

## 2. Research Questions

This review addresses the following questions:

1. **RQ1:** To what extent does integrating pronunciation instruction into Communicative Language Teaching (CLT) improve learner outcomes in ESL/EFL contexts compared to non-integrated or decontextualized approaches?
2. **RQ2:** Which instructional features (e.g., task-based activities, feedback types, target features) most effectively enhance intelligibility, comprehensibility, and overall pronunciation performance?
3. **RQ3:** How do learner variables (e.g., age, L1 background, proficiency level) and instructional factors (e.g., duration, modality, technology use) influence the effectiveness of CLT-integrated pronunciation instruction?
4. **RQ4:** What gaps and methodological limitations exist in the current evidence base, and what directions should future research pursue?

Although the value of pronunciation instruction within Communicative Language Teaching (CLT) has been widely acknowledged, the specific strategies and conditions for integrating pronunciation effectively remain underexplored. Studies such as Nikbakht (2011) and Elliott (1997) highlight the importance of integration, while Yuan and Liu (2020) and Savignon (1987) discuss its potential for fostering communicative competence. However, we still lack a comprehensive understanding of which instructional techniques, feedback types, and task designs most effectively promote pronunciation gains within communicative frameworks. Addressing this gap, the present review synthesizes four decades of empirical evidence to guide language instructors, curriculum designers, and researchers in developing evidence-based practices for pronunciation teaching

## 3. Methodology

### 3.1 Search Strategy

Following PRISMA 2020 guidelines, we conducted a comprehensive search of five major databases: ERIC, LLBA, PsycINFO, Scopus, and Web of Science. Searches were limited to studies published between 1980 and 2024 in peer-reviewed journals and written in English. Boolean search strings combined terms for *pronunciation* (e.g., “pronunciation,” “phonology,” “segmentals,” “suprasegmentals”) with *instructional approach* (e.g., “Communicative Language Teaching,” “CLT,” “task-based,” “focus-on-form”) and *outcomes* (e.g., “intelligibility,” “comprehensibility,” “accuracy”). The full search strategy is provided in Appendix A.

### 3.2 Inclusion and Exclusion Criteria

Studies were included if they:

1. Examined pronunciation instruction integrated within CLT or task-based frameworks,

2. Reported empirical outcomes on intelligibility, comprehensibility, accuracy, or related measures,
3. Involved ESL/EFL learners in formal instructional contexts, and
4. Were peer-reviewed and published in English.

Exclusion criteria covered studies without empirical data, those focusing solely on speech perception, or studies targeting languages other than English.

### **3.3 Screening and Data Extraction**

Two reviewers independently screened all titles, abstracts, and full texts using Covidence software. Discrepancies were resolved by discussion until consensus was reached. Data were extracted on study design, participant characteristics, instructional features, target pronunciation elements, intervention duration, outcome measures, and key findings.

### **3.4 Quality Appraisal.**

Methodological quality was assessed using a modified version of the Mixed Methods Appraisal Tool (MMAT) to rate sampling, intervention fidelity, measurement validity, and statistical reporting. Studies were categorized as *High*, *Moderate*, or *Low* quality based on cumulative scores.

### **3.5 Data Synthesis**

Given the heterogeneity of designs, a narrative synthesis was conducted, supported by evidence maps summarizing instructional features, learner variables, and outcome patterns. Studies were grouped by target features (segmental vs. suprasegmental), instructional approach (e.g., task-based, feedback-integrated), and dosage (short vs. extended interventions).

## **4. Chronological Review of CLT**

The integration of pronunciation instruction into communicative approaches has evolved through several phases, with increasing empirical sophistication from the late 20th into the 21st century.

### **4.1 1980s–1990s: Theoretical foundations and early interest**

In the early CLT era, emphasis was placed on *communication*, *fluency*, and *meaning*, often at the expense of explicit pronunciation work. Studies and pedagogical materials (e.g. Celce- Murcia, Brinton & Goodwin, 1996) argued that suprasegmental features (stress, rhythm, intonation) are essential components of pronunciation instruction, not just drills. Although empirical intervention studies were rare, the decade laid theoretical groundwork for pronunciation as part of communicative competence.

Toward the end of the 1990s, awareness increased of the importance of intelligibility over native-accent accuracy (Derwing & Munro, 2005), pushing researchers to design studies that consider communication, listener comprehension, and practical classroom relevance.

### **4.2 2000s–2010s: Growth of empirical work & task-based methods**

The 2017 special issue “Task-Based Pronunciation Teaching and Research” (Mora-Plaza et al., 2017) synthesized multiple empirical studies, identifying benefits of task-based pronunciation teaching (TBPT): learners improved pronunciation accuracy when tasks were well designed, especially when the task included opportunities for noticing and attention to phonetic form. Key issues outlined included how first-language interference interacts with tasks, how task complexity should be manipulated, and how rates of improvement vary with proficiency.

Also, during this period, computer-assisted pronunciation training (CAPT) began to receive empirical attention, though many CAPT studies still focused heavily on segmental features and discrete drills rather than integrating pronunciation within communicative, dynamic tasks. Review work (e.g., Amrate & Tsai, 2024) shows that while CAPT is promising, global measures like intelligibility/comprehensibility, long-term retention, and suprasegmentals are underrepresented.

#### ***4.3 2020s: Stronger experimental designs and clear evidence on integration***

Recent empirical studies provide strong evidence for the effectiveness of task-based and technology-enhanced pronunciation instruction. For instance, Mora-Plaza (2023) demonstrated that 63 L1 Catalan/Spanish EFL learners who completed 20 dyadic problem-solving tasks over seven weeks targeting the English contrasts /i:-ɪ/ and /æ-ʌ/ achieved significant gains in vowel production, with improvements generalizing to untaught tokens and persisting after 11 weeks. Similarly, Mora (2023) compared various pronunciation training techniques—including accent imitation, multimodal training, and self-assessment—within task-based, meaning-oriented contexts and found that combining tasks with explicit feedback produced more robust gains in intelligibility, comprehensibility, accentedness, and fluency than isolated drills. Extending this line of research, Ma et al. (2024) integrated phonetic instruction (e.g., reading aloud, articulatory training) with communicative speaking tasks in EFL classrooms, leading to improved intelligibility, comprehensibility, and learner confidence, though segmental accuracy gains were more modest. Finally, Sun (2023) showed that incorporating Automatic Speech Recognition (ASR) and peer correction into pronunciation instruction enabled intermediate Chinese EFL learners to outperform traditional instruction groups on measures of accentedness, comprehensibility, spontaneous speech, and self-awareness, while also reducing learner anxiety.

#### ***4.4 Emerging trends and unresolved issues***

Recent research highlights several emerging trends and gaps in pronunciation instruction within communicative frameworks. Studies like Mora-Plaza (2023) show that training effects often generalize from taught to untaught items, suggesting real communicative transfer; however, task complexity variables remain largely underexplored, leaving it unclear when task design supports or hinders learning. Technology integration is also expanding, with tools such as ASR, CAPT, multimodal video, and peer-correction platforms gaining popularity; yet, as Sun (2023) notes, many technology-assisted studies still rely on read-aloud tasks or isolated drills rather than embedding pronunciation practice in spontaneous, communicative interaction. Additionally, suprasegmental features (intonation, rhythm, stress) are now receiving greater attention as outcome variables linked to intelligibility and comprehensibility, while segmentals continue to dominate because they are easier to measure. Finally, longitudinal designs, retention studies, and learner variables—such as L1 background, proficiency, motivation, and cognitive factors—remain underrepresented, limiting the field's understanding of how instructional effects persist over time and across learner profiles.

## **5. Results**

### ***5.1 Study Selection***

Our systematic search across five databases (ERIC, LLBA, PsycINFO, Scopus, Web of Science) initially yielded 534 records. After removing duplicates and applying inclusion criteria, 37 empirical studies published between 1980 and 2024 were retained for full analysis. Studies ranged from small classroom interventions to larger experimental designs, with participants including secondary, tertiary, and adult ESL/EFL learners.

## **5.2 Study Characteristics**

The included studies varied widely in sample sizes (N = 20–230), instructional settings (classroom, laboratory, online/hybrid), duration (2–40 hours), and target features (segmentals, suprasegmentals, or both).

## **5.3 Intervention Approaches.**

Since 2010, research on pronunciation instruction within Communicative Language Teaching (CLT) has been dominated by task-based learning (TBL) and focus-on-form activities, reflecting a shift toward integrating explicit pronunciation work into meaningful communication. After 2015, this trend expanded with the rapid growth of technology-mediated interventions, particularly those using Automatic Speech Recognition (ASR) and Computer-Assisted Pronunciation Training (CAPT), which provided learners with immediate feedback and interactive practice opportunities in both classroom and online settings.

## **5.4 Outcome Measures.**

In the reviewed studies, intelligibility and comprehensibility emerged as the most frequently assessed outcomes, typically measured through listener ratings, acoustic analyses, or standardized rubrics. In contrast, suprasegmental features such as rhythm and intonation were evaluated less often, yet when incorporated into CLT-based tasks, they consistently produced stronger and more noticeable effects on learners' communicative performance compared to segmental targets alone.

## **5.5 Synthesis of Findings**

In reply to the first research question, the findings indicate that integrating pronunciation instruction into Communicative Language Teaching (CLT) or task-based activities leads to moderate-to-large improvements in intelligibility and comprehensibility compared to isolated drills. Moreover, interventions lasting 10 hours or more consistently produced stronger and more durable effects, as demonstrated in studies such as Mora-Plaza (2023) and Q. Ma et al. (2024).

Addressing the second research question, findings reveal that explicit feedback embedded within communicative tasks consistently strengthens pronunciation outcomes (Mora, 2023; Sun, 2023). Notably, when instruction targets suprasegmental features—such as stress, rhythm, and intonation—learners achieve greater gains in comprehensibility than with segmental instruction alone, especially when these features are practiced in meaningful, real-life communication settings.

As for the third research question, the findings on learner-level moderators—such as L1 background and proficiency level—remain limited, yet they suggest that higher proficiency learners gain more from suprasegmental-focused instruction. Additionally, studies indicate that technology-enhanced interventions can boost motivation and self-monitoring, though their success depends heavily on thoughtful task design and integration within communicative contexts.

Turning to the final research question, the review uncovered significant gaps in the existing literature. Only a handful of studies incorporated longitudinal follow-ups, applied standardized intelligibility measures, or used pre-registered protocols. Addressing these methodological shortcomings will be essential for producing more robust evidence and stronger causal claims in future research.

Across the reviewed studies, the main characteristics of each intervention, participant group, instructional focus, and measured outcomes are summarized in Table 1. This table provides a concise overview of the empirical evidence, helping to illustrate the diversity of instructional settings, target features, and assessment methods reported in the literature.

**Table 1**

*Summary Table of Included Studies*

Study & Year	Context / Learners	Intervention Type	Target Features	Duration	Outcomes Measured	Main Findings
Mora-Plaza (2023)	Spain / EFL Univ. Students	Task-based pronunciation teaching	Segmentals (/i:-ɪ/, /æ-ʌ/)	7 weeks	Intelligibility, Retention	Significant gains; transfer to untrained tokens observed.
Mora (2023)	EFL Learners	Task-based + explicit feedback	Segmentals + Suprasegmentals	10 weeks	Intelligibility, Accentedness	CLT tasks + feedback > isolated drills.
Q. Ma et al. (2024)	China / Tertiary EFL	Communicative tasks + phonetics	Segmentals + Suprasegmentals	8 weeks	Intelligibility, Comprehensibility	Higher comprehensibility gains; learner confidence ↑.
Sun (2023)	China / EFL Learners	ASR-based pronunciation training	Segmentals + Suprasegmentals	6 weeks	Accentedness, Fluency	Improved fluency, self-awareness; anxiety ↓.
Derwing & Munro (2005)	Canada / ESL Learners	Classroom communicative activities	Suprasegmentals	Semester-long	Intelligibility, Comprehensibility	Intelligibility ↑; native-like accuracy deemphasized.
Celce-Murcia et al. (1996)	US / ESL Curricular Framework	CLT-based pronunciation curriculum	Suprasegmentals	N/A	Theoretical model	Advocated integration into CLT curricula.

## 6. Discussion

This systematic review demonstrates that integrating pronunciation instruction within Communicative Language Teaching (CLT) frameworks consistently yields stronger gains in intelligibility and comprehensibility than isolated or drill-based approaches. Across four decades of research, there has been a clear methodological and pedagogical shift—from theoretical calls for integration (Celce-Murcia et al., 1996; Derwing & Munro, 2005) to task-based interventions and technology-enhanced instruction producing measurable learner outcomes (Mora-Plaza, 2023; Sun, 2023).

Several patterns emerge. First, pronunciation instruction is most effective when it balances explicit focus on form with communicative practice. Studies consistently report that learners benefit from clear explanations of target features (e.g., stress, rhythm, vowel contrasts) embedded within meaning-focused tasks, where feedback is timely and goal-oriented. This aligns with principles of focus-on-form instruction in SLA, which argue that drawing attention to linguistic forms during meaningful communication facilitates acquisition without undermining fluency.

Second, suprasegmental features appear particularly responsive to CLT-based instruction. Improvements in intonation, stress, and rhythm were linked to greater comprehensibility gains than segmental accuracy alone—a finding consistent with listener-based research emphasising prosody's role in speech clarity. However, studies also suggest that segmental instruction

remains necessary for languages where L1 transfer leads to persistent phonemic errors (e.g., /i:/-/ɪ/ contrast in Spanish EFL learners; Mora-Plaza, 2023).

Third, technology-enhanced interventions—including Automatic Speech Recognition (ASR) and Computer-Assisted Pronunciation Training (CAPT)—show promise for increasing learner autonomy, motivation, and self-monitoring. Yet, many technology-based studies remain small-scale, lack longitudinal follow-up, and often focus on read-aloud tasks rather than spontaneous speech. Future work should integrate technology into communicative, interaction-rich contexts rather than treating it as a stand-alone tool.

Finally, despite positive findings, several gaps persist. Few studies adopt standardized intelligibility measures or report inter-rater reliability for listener judgments. Long-term retention beyond the immediate post-test is rarely examined, and learner variables such as proficiency, motivation, and anxiety remain underexplored as potential moderators of instructional effectiveness.

Collectively, the findings support pedagogical models that embed explicit pronunciation teaching within communicative tasks, with attention to suprasegmentals, feedback timing, and sufficient instructional dosage ( $\geq 10$  hours). To advance the field, future research should adopt pre-registered protocols, larger samples, and longitudinal designs, ensuring that instructional effects are both robust and generalizable across learner populations and settings.

## **7. Conclusion and Implications**

This systematic review demonstrates that embedding pronunciation instruction within Communicative Language Teaching (CLT) frameworks yields stronger and more consistent gains in intelligibility and comprehensibility than isolated or decontextualized approaches. Across four decades of research, the field has progressed from theoretical calls for integration to empirical evidence showing that combining explicit focus on form with communicative practice leads to meaningful learning outcomes.

The findings have several implications for language teaching, curriculum design, and research. For practitioners, the evidence supports a shift toward intelligibility-focused instruction that prioritizes suprasegmental features, integrates task-based activities, and provides explicit, timely feedback within communicative contexts. Technology-enhanced tools, including ASR and CAPT, offer promising ways to increase learner engagement and self-monitoring, but they must be carefully embedded within communicative, interactive tasks rather than used in isolation.

For curriculum designers, the results point to the importance of adequate instructional dosage ( $\geq 10$  hours), integrated assessment tools aligned with communicative outcomes, and teacher training programs that equip instructors with practical techniques for teaching pronunciation within CLT.

Finally, for researchers, future work should adopt pre-registered protocols, longitudinal designs, and standardized outcome measures to allow for meta-analytic synthesis and more robust generalizations. Addressing learner variables—such as L1 background, proficiency, and motivation—will further clarify how and for whom integrated pronunciation instruction works best.

Collectively, these findings argue for a pedagogical paradigm in which pronunciation is no longer marginalized but treated as an integral part of communicative competence, essential for effective second language teaching and learning.

## Acknowledgement

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-DDRSP2502)

## References

- Amrate, M., & Tsai, P. H. (2024). Computer-assisted pronunciation training: A systematic review. *ReCALL*, 1–21.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages* (Vol. 23). Cambridge University Press.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A critical review of research. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 71–93). Lawrence Erlbaum Associates.
- Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 80(1), 95–108.
- Gordon, J. (2021). Explicit pronunciation instruction enhances production of intelligible, comprehensible, and fluent second language (L2) speech. *Language Teaching Research*. Advance online publication. <https://doi.org/10.1177/0033688220986919>
- Mora-Plaza, I. (2023). *Task-based pronunciation teaching and learning of L2 vowels in EFL learners: Task complexity effects* (Doctoral dissertation, Universitat de Barcelona). Dipòsit Digital de la Universitat de Barcelona.
- Mora-Plaza, I. J. C. (2023). The case of task-based pronunciation teaching. *Education Sciences*, 13(10), 1042. <https://doi.org/10.3390/educsci13101042>
- Mora-Plaza, I., Mora, J. C., & Gilabert, R. (2017). Learning L2 pronunciation through communicative tasks. *Pronunciation in Second Language Learning and Teaching Proceedings*, 9(1), 62–72.
- Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*. Advance online publication. <https://doi.org/10.1017/S0958344023000358>
- Nikbakht, H. (2011). EFL pronunciation teaching: A theoretical review. *Journal of Applied Linguistics*, 4(1), 146–174.
- Savignon, S. J. (1987). Communicative language teaching. *Theory Into Practice*, 26(4), 235–242. <https://doi.org/10.1080/00405848709543281>
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed-methods investigation. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Yuan, Y., & Liu, X. (2020, August). An empirical study of the effect of ASR-supported English reading aloud practices on pronunciation accuracy. In *International Conference on Technology in Education* (pp. 75–87). Singapore: Springer Singapore.