CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Teaching Connected Speech through Spiral Pedagogy: Addressing ESL Listening Barriers at the Tertiary Level

by

Bushra Hassan Mirza

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Management & Social Sciences

Department of English

2026

CERTIFICATE OF APPROVAL

Teaching Connected Speech through Spiral Pedagogy: Addressing ESL Listening Barriers at the Tertiary Level

by

Bushra Hassan Mirza

(MEN243006)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Hazrat Umar	NUML, Islamabad
(b)	Internal Examiner	Dr. M. Saqib Zaigham	CUST, Islamabad

Dr. Muhammad Umar Farooq
Thesis Supervisor
May, 2026

Dr. Muhammad Umar Farooq
Head
Department of English
May, 2026

Dr. Arshad Hassan
Dean
Faculty of Management & Social Sci.
May, 2026

Author's Declaration

I, **Bushra Hassan Mirza** hereby state that my MS thesis titled "**Teaching Connected Speech through Spiral Pedagogy: Addressing ESL Listening Barriers at the Tertiary Level**" is my original work and has not been submitted, in whole or in part, for the award of any degree at Capital University of Science and Technology, Islamabad, or at any other institution within Pakistan or abroad. I further affirm that, should any part of this declaration be found to be false at any stage, including after the award of the degree, the University reserves the right to withdraw my MS degree in accordance with its rules and regulations.

(Bushra Hassan Mirza)

Registration No: MEN243006

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Teaching Connected Speech through Spiral Pedagogy: Addressing ESL Listening Barriers at the Tertiary Level**" is entirely my own work and that no substantial contribution has been made by any other individual. Any assistance or support received during the course of this research has been appropriately acknowledged, and the thesis has been written solely by me.

I fully understand and acknowledge the zero-tolerance policy of the Higher Education Commission (HEC) of Pakistan and Capital University of Science and Technology regarding plagiarism. I hereby declare that no part of this thesis has been plagiarized and that all sources of information, data, or ideas used in this research have been properly cited and referenced in accordance with academic standards.

I further undertake that if I am found guilty of any form of plagiarism in this thesis at any stage, including after the award of the MS degree, the university reserves the right to withdraw or revoke my degree. I also acknowledge that HEC and the university may publish my name on their official websites among the list of students found to have submitted plagiarized work.

(Bushra Hassan Mirza)

Registration No: MEN243006

Acknowledgement

This thesis is dedicated to my husband, for his patience, faith, and constant encouragement; to my children, for their understanding, resilience, and belief in my purpose; and to my parents, for their unwavering support, guidance, and prayers. Their collective strength made this journey possible.

I extend my sincere gratitude to my supervisor, Dr. Muhammad Umar Farooq, for his invaluable guidance, insightful feedback, and steadfast support throughout this research. His expertise played a crucial role in shaping this study.

I am also thankful to the faculty of the Department for their academic support, to the participants for their time and cooperation, and to Jess Thida for generously providing pro bono voice recordings for the pre- and post-tests.

(Bushra Hassan Mirza)

Abstract

Out of all the four language skills, listening comprehension is widely recognized as one of the most difficult skills for Pakistani learners of English. Although students are regularly exposed to English through academic settings and digital media, many still struggle to understand fluent spoken language. This difficulty often stems from problems in recognizing connected speech, where sounds are reduced, blended, or modified in natural communication. As a result, learners experience decoding failure at the level of phonological perception rather than vocabulary or grammar. The present study explores whether teaching connected speech through a spiral pedagogy can help reduce these listening barriers at the tertiary level. A quasi-experimental design was adopted, involving a control group and an experimental group. Five connected speech features; assimilation, elision, flapping, weak forms, and contractions were systematically taught and assessed. Researcher-developed pre-tests and post-tests were used to measure learners' perception of these features. Because the test sections contained unequal numbers of items, weighted scoring was applied to calculate individual listening scores, while aspect-wise averages were used for group-level analysis. Quantitative data were analyzed using descriptive statistics and paired-samples t-tests, supported by qualitative classroom observations. The results show that learners in the experimental group made noticeable gains across all connected speech features, particularly in elision, flapping, and weak forms. In contrast, the control group demonstrated only minor improvement, which appears to reflect natural exposure rather than instructional impact. Analysis of individual gain scores further revealed considerable variation in learner progress, highlighting the role of explicit and repeated exposure to connected speech features. Overall, the findings suggest that a spiral-based approach to teaching connected speech can significantly improve learners' ability to process fluent spoken English. The study highlights the importance of addressing phonological aspects of listening and offers practical implications for ESL instruction, curriculum development, and teacher training in Pakistani higher education context.

Contents

List of Figures

List of Tables

List of Abbreviations

CELTA Certificate in Teaching English to Adults

CLT Cognitive Load Theory

EFL English as a Foreign Language

ESL English as a Second Language

HEC Higher Education Commission

IPA International Phonetic Alphabet

L2 Second Language

TP Teaching Practice of 3 hours

Chapter 1

Introduction

While Pakistani English is a recognized variety of English with its own phonological system (Rahman, 1990) (Baumgardner, 1993), this does not eliminate the need for learners to process connected speech in other widely encountered varieties, particularly American English, which dominates global academic and media contexts. Having taught English at the university level for over a decade, I had observed that many students struggle to understand fluent, natural speech, even after considerable exposure to English as a second language. This recurring challenge motivated the present study, which seeks to identify the connected speech features that Pakistani learners of English find most difficult. The study also proposes a focused instructional approach and examines its effectiveness in helping learners improve their perception of the American accent they most frequently encounter through social media.

While the World Englishes framework rightly recognizes many varieties of English as legitimate, the focus of the present study is not on accent standardization or speech production, but on listening comprehension in contexts where learners are exposed to fluent, naturally connected speech. In academic lectures, digital media, and online learning platforms, Pakistani learners frequently encounter reduced and rapid speech forms such as weak forms, elision, and assimilation, particularly in widely circulating international varieties of English. Difficulty in perceiving these features creates a practical barrier to comprehension, regardless of the legitimacy of local varieties. Consequently, the decision to focus on connected speech aspects

in this study was pedagogical rather than ideological, aimed at equipping learners with perceptual strategies to decode authentic spoken input across contexts, rather than promoting or privileging any single model of English.

1.1 Background of the Study

Despite English being taught in Pakistan from the early years of schooling and practiced extensively across successive academic levels, only a limited proportion of learners achieve communicative spoken proficiency, and even then, mostly it is achieved through self-practice. If we analyze the reasons behind it, what we come to an understanding that, there were little to none trainings for the teachers, the focus was on reading and writing alone, spoken English was never prioritized, and listening was totally ignored. It is a known fact that output requires input, thus when we talk about speaking English, there is a lack of individuals who can speak correct and fluent English because they were never involved in active listening tasks. We know that reading builds writing skills and listening leads to speaking skills. English teachers ask students to immerse themselves in language by watching movies or listen to English songs, yet no one teaches how to listen. Thus, we need teachers who can reinforce speaking through listening activities. Listening modules have to be indispensable for teaching English as a spoken language as it serves as a foundation on which speaking proficiency is constructed. Spoken language has suprasegmental features and connected speech aspects that makes it different from the text. Therefore, it is of utmost importance that these be taught in order for the speakers to sound natural and intelligible. Besides, students who do not get to develop speaking skills face discrimination in university regardless of their knowledge, marks and intelligence which leads to lack of confidence, loneliness and low self-esteem. In order for the learners to be able to internalize the suprasegmental features of any language through listening practices, it is important to develop their ears. It is important to reiterate that the aim of this study is not to promote the learning of any particular accent, but to enable learners to comprehend naturally spoken English for meaningful learning and effective communication in competitive academic and professional contexts.

The benefits of developing the ear are twofold; it will let the learners assimilate more content through increased intelligibility of the content, and it will help internalize the suprasegmental features and connected speech aspects on subconscious level which in turn will result in learners exhibiting more natural sounding speech patterns. Research suggests that individual differences in auditory processing ability, such as having a "good ear," play significant roles in learners' ability to perceive and benefit from spoken input in second language contexts (Saito, 2023).

1.2 Connected Speech and Listening Comprehension

Connected speech refers to the natural flow of spoken language in which words are not produced in isolation but are shaped by their phonological environment, ease of articulation and speech rate. In fluent speech, speakers routinely modify words or word boundaries through processes such as assimilation, elision, reduction, contractions, weak forms, and flapping. These processes contribute to speech efficiency and fluency but simultaneously increase perceptual difficulty for second language (L2) listeners, particularly those accustomed to slow, carefully articulated classroom input (Brown and Hilferty, 2006) (Celce-Murcia et al., 2010). In countries such as Pakistan where students mostly learn English by reading under untrained teachers, they do not come across natural speech and are accustomed to pronouncing the words with stress on every sound and syllable because this is how their mother tongue Urdu is spoken. When acoustic input exceeds processing capacity, learners rely on linguistic expectations set by mother tongue, resulting in perceived but non-existent forms. Thus, the influence of L1 on L2 is quite strong and they find the natural reduced speech frustrating. To cater to this problem, the module is designed following a bottom-up approach.

Listening comprehension in a second language involves the interaction of bottom-up and top-down processing. While top-down processes rely on contextual knowledge, expectations, and prior experience, bottom-up processing requires accurate decoding of the acoustic signal into meaningful linguistic units (Field, 2008). Research

consistently indicates that many L2 listening problems originate at the bottom-up level, where learners fail to segment the speech stream due to altered phonological forms in connected speech (Rost, 2011). When listeners cannot identify where one word ends and another begins, comprehension breakdowns occur even when vocabulary and grammar knowledge are sufficient.

Connected speech poses a particular challenge because it disrupts the one-to-one correspondence between written and spoken language. Learners often approach listening with expectations shaped by orthographic representations, assuming that spoken input will reflect textbook forms. However, in authentic speech, function words are frequently reduced, consonants may be deleted or assimilated, and sounds may merge across word boundaries (Cauldwell, 2013). As a result, learners may fail to recognize familiar words when they appear in reduced or contracted forms, leading to misinterpretation or complete loss of meaning.

Empirical studies in L2 listening have demonstrated that lack of awareness of connected speech features significantly affects comprehension accuracy. Field (2008) argues that learners often "hear what they expect" rather than what is acoustically present, causing them to mis-segment input and construct inaccurate meanings. Similarly, Brown and Hilferty (2006) note that fast speech combined with phonological reduction increases processing demands, overwhelming learners' working memory and hindering real-time comprehension. These difficulties are exacerbated in examination-oriented contexts where listening instruction prioritizes test strategies over perceptual training.

Another key issue is the rate of speech. In natural interaction, speech is produced at a pace that leaves little time for conscious decoding. Native speakers rely on automatized perceptual skills to process reduced forms efficiently, whereas L2 learners often attempt word-by-word processing, which is incompatible with the speed of authentic input (Rost, 2011). Without prior exposure to connected speech patterns, learners struggle to keep pace with incoming information, resulting in increased anxiety and reduced confidence in listening tasks.

Learner perceptions further contribute to the complexity of connected speech processing. Studies indicate that learners frequently attribute listening failure to

unknown vocabulary or speaker accent, rather than to phonological modification in connected speech (Goh, 2000) (Cauldwell, 2013). This misattribution discourages learners from focusing on phonological form and limits their ability to develop effective listening strategies. Many local researches suggest including listening in ESL programs but they do not answer how. Consequently, repeated exposure without explicit guidance may reinforce listening frustration rather than improve comprehension.

Overall, the literature suggests a strong and direct relationship between connected speech and listening comprehension. Difficulties in recognizing reduced and modified forms interfere with word recognition, disrupt meaning construction, and place excessive demands on cognitive resources. These findings underscore the need for instructional approaches that address connected speech explicitly, enabling learners to develop perceptual awareness and improve bottom-up processing. Understanding the role of connected speech in listening comprehension is therefore essential for designing effective pedagogical interventions aimed at improving L2 listening proficiency.

1.3 Problem Statement

The consumption of English content on social media results in self-learning the language through top-down approach unconsciously, but do the viewers actually understand the spontaneous dialogues and fluent English, or do they rely on subtitles to understand the content? This was the first question that led to my interest in discovering how non-native speakers process the connected speech. To develop speaking skills in a country which tops in having freelancers who are mostly getting work from west and have to communicate in English, it is important that their listening skills are enhanced which will automatically boost their speaking skills by helping internalize the connected speech aspects.

Classroom observations in a tertiary-level ESL context further highlighted the nature of the listening difficulties addressed in this study. Despite demonstrating adequate proficiency in reading and writing, learners frequently struggled to follow

naturally spoken English, particularly when speech was delivered at normal or fast pace. During listening sessions such as video lectures and interviews, many students appeared to rely on isolated content words to infer meaning, while reduced function words and connected speech features often went unnoticed. This tendency led to partial or inaccurate interpretation of spoken input, suggesting that the underlying difficulty was not lexical knowledge but limited sensitivity to phonological features in continuous speech. These observations reinforced the need to examine connected speech perception as a specific source of listening difficulty.

1.4 Research Gap

Local Pakistani studies consistently acknowledge the importance of listening as a core skill in ESL classrooms and recommend its inclusion in language instruction e.g., (Nawaz, 2021) (Hussain and Javed, 2022) (Riaz et al., 2023). However, these studies largely remain diagnostic in nature, focusing on learners' difficulties, strategy use, or contextual constraints, rather than offering pedagogically grounded models of listening instruction. In particular, they do not address how listening should be systematically taught, what specific features of spoken English should be targeted, or how learners' perceptual abilities can be progressively developed.

More importantly, despite repeated acknowledgment that Pakistani ESL learners struggle with fast, natural speech, no empirical study to date has examined the explicit teaching of connected speech features such as reduction, assimilation, and alveolar flapping as a means of overcoming connected speech processing difficulty in the Pakistani ESL context. Existing research attributes listening problems to factors such as lack of exposure to authentic input, L1 interference, and pedagogical neglect e.g.,(Nawaz, 2021), yet stops short of experimentally investigating instructional interventions that target these issues at the perceptual level.

The present study seeks to fill this gap by investigating the teaching of aspects of connected speech through a spiral methodology, in which listening input and tasks are recycled and gradually expanded in complexity to build listening comprehension incrementally. By examining how Pakistani ESL learners perceive and

process connected speech features when instruction is structured spirally, this study moves beyond identifying listening problems to empirically exploring a principled instructional response to connected speech processing difficulty in the Pakistani ESL classroom.

1.5 Theoretical Orientation of the Study

The present study is grounded in cognitive and psycholinguistic theories that explain how second language listeners process spoken input and why difficulties arise when speech is produced in its natural, connected form. Specifically, the study draws on Cognitive Load Theory, speech perception theories, and models of top-down and bottom-up processing to provide a theoretical basis for investigating the role of connected speech instruction in improving listening comprehension.

Together, these theoretical perspectives suggest that listening comprehension difficulties in L2 contexts are not solely cognitive or strategic but are closely tied to perceptual challenges at the phonological level. Cognitive Load Theory explains how connected speech increases processing demands, speech perception theories account for segmentation and recognition difficulties, and processing models highlight the need for balanced top-down and bottom-up interaction.

This theoretical orientation supports the pedagogical assumption underlying the present study: that explicit instruction in connected speech aspects can reduce cognitive load, enhance perceptual accuracy, and ultimately improve listening comprehension.

1.6 Spiral Pedagogy in Contemporary Listening Instruction

Listening comprehension in a second language develops gradually and requires repeated exposure to linguistic features across varied contexts. In response to this, the present study adopts a spiral-based approach to listening instruction, in which key connected speech features are introduced incrementally and systematically

revisited with increasing complexity. While the spiral curriculum was initially conceptualized in earlier educational theory, its underlying principles remain highly relevant in contemporary second language pedagogy.

Current cognitive and usage-based perspectives emphasize the importance of managing processing demands, recycling input, and allowing learners to form perceptual categories through repeated encounters with meaningful language. Within this framework, spiral sequencing supports listening development by preventing cognitive overload and enabling learners to gradually attune to reduced and modified speech forms commonly found in authentic spoken English.

Accordingly, the instructional design of this study reflects a modern, cognitively informed interpretation of spiral pedagogy applied to ESL listening at the tertiary level.

1.7 Research Objectives

Objectives of the study are as follows.

i. To Propose a practical solution; an inductive spiral-based teaching methodology for developing ears which can be implemented to all grades and levels.

ii. Analyze whether the proposed methodology lowers cognitive overload in ESL learners.

1.8 Research Questions

This research will answer the following questions.

i. To what extent does a spiral-based teaching module enhance ESL learners' processing of fluent General American English speech?

ii. To what extent does the spiral-based methodology reduce connected speech processing difficulty among ESL learners?

1.9 Significance of the Study

This research will help administrators, teachers, and academia understand the importance of teaching listening and that including listening in ESL teaching is not challenging. This research provides a roadmap in the form of a module that will help students academically by enabling the students to comprehend the foreign content; academic lectures and clips on websites like YouTube, making them understand the language used in movies and serials on Netflix and other mediums. This will further strengthen their speech by helping internalize other aspects of English like, plosives as well as supra-segmental features like, stress, intonation, and so on. The study proposes a practical, multi-level instructional approach; the Spiral Remedy designed to strengthen listening skills through progressive, scaffolded exposure to connected speech patterns at university level.

1.10 Delimitations of the Study

Although the proposed instructional approach is potentially applicable across different educational levels, the present study was limited to university students enrolled in Communication Studies who demonstrated a reasonable level of proficiency in English reading and writing. As a result, the findings may not be directly generalizable to learners with minimal English proficiency.

The study was conducted with a relatively small number of participants from a single institution. As a result, the findings may not be fully generalizable to all Pakistani learners of English or to learners in different educational contexts.

In addition, while several phonological features contribute to accurate pronunciation and listening comprehension such as word stress, sentence stress, intonation, plosives, and phonemic contrasts between English and Urdu sounds e.g., **/d/, /t/, /r/**, and **/w/**, the study focused exclusively on selected aspects of connected speech. These included contractions, weak forms, elision, assimilation, and alveolar flapping, the latter being more characteristic of American English. Other phonological dimensions were therefore not examined.

Furthermore, as American English predominates in instructional materials and media exposure in the local context, the study was restricted to the American accent. Consequently, the findings may not fully reflect learners' processing of other English varieties, such as British or Australian accents.

Finally, due to time constraints, the instructional intervention was limited to 32 contact hours. A longer intervention period must have yielded deeper and more sustained gains in learners' perception of connected speech features.

Chapter 2

Literature Review

2.1 Introduction

Listening comprehension is a core component of second language proficiency; however, it remains one of the most challenging skills for L2 learners to develop effectively (Vandergrift and Goh, 2012). In many ESL contexts, listening instruction relies heavily on exposure and strategy training, with limited emphasis on the phonological characteristics of natural spoken language (Field, 2008). As a result, learners often experience difficulty in processing authentic speech, particularly when spoken input contains reduced and connected forms that differ substantially from citation forms typically encountered in instructional materials.

Research in speech perception and listening comprehension indicates that successful listening requires accurate decoding of the acoustic signal in addition to higher-level inferencing and meaning construction (Rost, 2011). Features of connected speech, such as assimilation, elision, linking, weak forms, and flapping, frequently obscure word boundaries and alter surface forms, increasing perceptual difficulty for L2 listeners (Brown and Hilferty, 2006) (Celce-Murcia et al., 2010). When learners lack awareness of these features, comprehension breakdowns are likely to occur despite adequate lexical or grammatical knowledge.

The literature reviewed in this chapter is organized thematically to trace the relationship between listening processes, connected speech phenomena, and instructional practices. It first examines theoretical models of listening and speech

perception, followed by a review of research on connected speech and its impact on L2 listening difficulties. The chapter then analyses pedagogical approaches to listening instruction, with particular focus on studies that have investigated explicit instruction in connected speech features and their effects on listening comprehension (Field, 2008) (Cauldwell, 2013). Empirical findings are critically evaluated to identify methodological strengths, limitations, and areas of underrepresentation.

Finally, the review considers studies conducted in ESL contexts comparable to Pakistan, where listening instruction and phonological training remain limited within formal curricula (Mahmood, 2018) (Rahman and Khan, 2020). By synthesizing global and contextual research, this chapter identifies a clear gap in experimental studies that systematically examine the impact of teaching connected speech aspects on listening comprehension. This gap provides the theoretical and empirical justification for the present study and informs the research design outlined in the subsequent chapter.

2.2 Listening Comprehension: Nature, Types and Processes

Listening comprehension is a complex, multidimensional skill that involves the active construction of meaning from spoken input rather than the passive reception of sounds. In second language acquisition, listening requires learners to decode the acoustic signal, recognize linguistic units, and integrate incoming information with prior knowledge in real time (Rost, 2011)(Vandergrift and Goh, 2012). Unlike reading, listening offers limited opportunity for revision or control over the rate of input, which places substantial cognitive demands on the listener, particularly in L2 contexts.

Listening comprehension is a complex, multidimensional skill that involves the active construction of meaning from spoken input rather than the passive reception of sounds. In second language acquisition, listening requires learners to decode the acoustic signal, recognize linguistic units, and integrate incoming information with prior knowledge in real time (Rost, 2011)(Vandergrift and Goh, 2012)(Goh, 2000).

This process operates across multiple levels of listening, most notably discriminative and comprehensive listening. Discriminative listening refers to the ability to perceive and distinguish between phonological features such as sounds, stress, intonation, and connected speech, forming the foundation of accurate auditory processing (Brown and Hilferty, 2006) (Field, 2008). In contrast, comprehensive listening involves the higher-order processing of meaning, where listeners interpret and construct semantic understanding from the input. These two types are interdependent, as effective comprehension is contingent upon accurate lower-level discrimination of the speech signal. Unlike reading, listening offers limited opportunity for revision or control over the rate of input, which places substantial cognitive demands on the listener, particularly in L2 contexts. The need to simultaneously manage phonological decoding and meaning construction further intensifies cognitive load, making listening a particularly demanding skill for language learners (Vandergrift and Goh, 2012) (Goh, 2000) (Field, 2008).

2.2.1 Models of Listening Comprehension

Early models of listening comprehension emphasized bottom-up processing, in which comprehension is achieved through the linear decoding of sounds into phonemes, words, and syntactic structures (Anderson, 1995). From this perspective, successful listening depends largely on accurate perception of the speech signal. However, purely bottom-up models were later criticized for underestimating the role of the listener's background knowledge and expectations.

In contrast, top-down models view listening as a meaning-driven process, where listeners use contextual cues, prior knowledge, and schematic information to interpret spoken input (Rumelhart, 1980)(Rost, 2011). While this approach accounts for predictive processing and inferencing, it may overlook the importance of accurate phonological decoding, particularly for L2 learners with limited exposure to natural speech.

More recent research supports interactive models of listening, which propose that bottom-up and top-down processes operate simultaneously and dynamically during comprehension (Vandergrift and Goh, 2012). In these models, perceptual

input and cognitive expectations continuously interact, allowing listeners to adjust interpretations as new information becomes available. Such models are particularly relevant to L2 listening, as breakdowns often occur when bottom-up processing is compromised by unfamiliar phonological features, such as connected speech phenomena.

2.2.2 Cognitive Processes Involved in Listening

Listening comprehension engages multiple cognitive processes, including attention, working memory, and perceptual processing. As spoken input unfolds rapidly, listeners must allocate attentional resources efficiently while retaining segments of speech in working memory long enough to integrate them into meaningful units (Field, 2008). When the cognitive load exceeds the listener's processing capacity, comprehension is likely to deteriorate.

Research suggests that L2 listeners often expend excessive cognitive effort on low-level decoding, leaving fewer resources available for higher-level meaning construction (Goh, 2000) (Vandergrift and Goh, 2012). This imbalance is particularly evident when learners encounter reduced or altered phonological forms in connected speech. In such cases, the failure to recognize familiar words in their reduced forms disrupts processing and increases reliance on guessing or inferencing, which may not always be successful.

Furthermore, speech perception is influenced by learners' first language phonological systems, which shape how acoustic input is categorized and interpreted (Cutler, 2012). When L2 input deviates from learners' expectations, as is often the case in fast or connected speech, perceptual confusion and misinterpretation are likely to occur.

2.2.3 Difficulties in Second Language Listening Comprehension

Empirical studies consistently identify listening as one of the most problematic skills for L2 learners (Graham, 2006) (Field, 2008). Common difficulties include

inability to segment the speech stream, failure to recognize known words, and difficulty coping with speed and variability in natural speech. These problems are frequently attributed to limited exposure to authentic spoken input and insufficient instruction in the phonological features of spoken language.

Connected speech poses a particular challenge, as it alters canonical word forms through processes such as assimilation, elision, and reduction. As a result, learners may perceive spoken input as a continuous, unintelligible stream rather than as discrete lexical items (Brown and Hilferty, 2006). This perceptual difficulty often leads learners to report that native speakers speak "too fast," when in reality the issue lies in phonological processing rather than speed alone.

In many ESL classrooms, listening instruction prioritizes comprehension questions and strategy use while neglecting the perceptual foundations of listening (Field, 2008) (Cauldwell, 2013). Consequently, learners are assessed on comprehension outcomes without being equipped with the phonological tools necessary to process authentic speech effectively. These limitations underscore the need for instructional approaches that address listening as a perceptual and cognitive process, rather than solely as a test-based skill.

2.3 Phonological Features of Connected Speech

Connected speech refers to the natural flow of spoken language in which words are not produced in isolation but as part of a continuous stream. In authentic spoken English, phonological forms frequently diverge from their citation forms due to articulatory efficiency, rhythmic patterns, and prosodic constraints (Brown and Hilferty, 2006) (Celce-Murcia et al., 2010). These variations are systematic and predictable, yet they often remain underrepresented in formal language instruction, contributing to listening difficulties for L2 learners.

In classroom contexts, learners are typically exposed to carefully articulated speech, whereas real-life listening involves reduced, modified, and compressed forms (Field, 2008). The following subsections outline key phonological features of connected speech that significantly affect listening comprehension.

2.3.1 Assimilation

Assimilation occurs when a sound changes to become more similar to a neighboring sound in terms of place, manner, or voicing of articulation. This process facilitates smooth and efficient speech production but alters the surface form of words (Roach, 2009). Assimilation may be regressive, progressive, or coalescent, depending on the direction of influence.

For example, in rapid speech, the phrase "good boy" may be realized as **gooboy**, where the alveolar /d/ assimilates to the bilabial /b/. Such changes can obscure word boundaries and hinder word recognition for L2 listeners. Studies show that learners unfamiliar with assimilation patterns often fail to recognize known lexical items, leading to comprehension breakdowns (Brown and Hilferty, 2006).

2.3.2 Elision

Elision refers to the omission of sounds, most commonly consonants or vowels, in connected speech. It frequently occurs in consonant clusters or unstressed syllables, particularly in rapid or informal speech (Roach, 2009). For instance, "next day" may be pronounced as **nex day**, with the /t/ omitted. While elision reduces articulatory effort, it poses a serious challenge for L2 listeners because the absence of expected sounds disrupts word recognition and segmentation. Learners may interpret elided forms as incomplete or unfamiliar words rather than as natural reductions, resulting in listening difficulty and loss of confidence (Field, 2008).

2.3.3 Contractions

Contractions involve the phonological and orthographic reduction of two words into a single, compressed form, most commonly through the omission of sounds and the use of unstressed vowels. Examples include "I am" → "I'm," "do not" → "don't," and "they have" → "they've" (Celce-Murcia et al., 2010). Contractions are highly frequent in spoken English, particularly in informal and semi-formal contexts.

Despite their frequency, contractions often present listening difficulties for L2 learners because they differ markedly from their full forms, which are typically emphasized in written instruction. Learners may fail to identify grammatical elements such as negation, tense, or modality when contractions are used, leading to misinterpretation of meaning (Field, 2008). Explicit instruction in contractions has been shown to improve learners' ability to recognize reduced grammatical forms and process spoken input more efficiently (Cauldwell, 2013).

2.3.4 Weak Forms

Weak forms occur when function words such as articles, prepositions, auxiliary verbs, and conjunctions are pronounced in a reduced manner in unstressed positions. For example, "can" may be realized with a schwa instead of /ae/ sound, and "to" with a schwa instead of /u/ sound (Roach, 2009). Weak forms are essential to the rhythmic structure of English and are pervasive in natural speech. L2 learners often struggle with weak forms because they are typically taught strong forms in isolation. Consequently, learners may fail to recognize highly frequent grammatical words during listening, which disrupts syntactic parsing and overall comprehension (Cauldwell, 2013). Research indicates that awareness-raising and focused listening practice can significantly enhance learners' perception of weak forms (Celce-Murcia et al., 2010).

2.3.5 Flapping

Flapping is a phonological process in which intervocalic /t/ and /d/ sounds are realized as a voiced alveolar flap /ɾ/, as in "better" and "ladder" (Roach, 2009). Although it is a common phenomenon across English accents, including the British accent, it is particularly common in American English. This process neutralizes the contrast between /t/ and /d/ in specific phonological environments, which can be confusing for L2 listeners. Learning it assists in the comprehension of different accents as well.

For learners whose first languages do not include an equivalent sound, flapping may result in misperception or lexical ambiguity. Learners may interpret flapped sounds

as /r/ or fail to recognize the word entirely, especially in fast speech (Celce-Murcia et al., 2010). Instruction that explicitly addresses flapping can help learners develop more accurate perceptual representations of spoken English.

2.3.6 Implications of Connected Speech Features for Listening Comprehension

The phonological features discussed above contribute to the fluency and naturalness of spoken English but also increase perceptual complexity for L2 learners. When learners lack awareness of these features, they may rely excessively on top-down guessing strategies, which often results in inaccurate comprehension (Field, 2008). Increasing evidence suggests that explicit instruction in connected speech features enhances learners' perceptual sensitivity and supports more efficient processing of authentic spoken input (Cauldwell, 2013).

Understanding connected speech as a core component of listening comprehension rather than merely a pronunciation concern provides a strong rationale for instructional interventions. This perspective underpins the present study and informs the pedagogical focus examined in subsequent sections of this literature review.

2.4 Listening Difficulties Caused by Connected Speech

Listening difficulties in second language contexts are frequently attributed to limited vocabulary or insufficient grammatical knowledge; however, research increasingly indicates that many comprehension problems arise from perceptual challenges associated with connected speech (Field, 2008)(Rost, 2011). people speak naturally, When L2 learners hear natural speech rather than carefully pronounced citation forms, they struggle to understand the speech stream, even when the words are familiar to them. This section examines the major listening difficulties caused by connected speech, with particular emphasis on word boundary confusion, reduced forms, fast speech, and learner perceptions of spoken English.

2.4.1 Word Boundary Confusion

One of the most commonly reported difficulties in L2 listening is the inability to segment continuous speech into meaningful units. In natural spoken English, words are not separated by pauses; instead, they are produced as part of a continuous acoustic stream. Connected speech features such as assimilation, elision, weak forms, and contractions obscure clear word boundaries, making segmentation particularly challenging for L2 listeners (Field, 2008).

Learners who only rely on written representations of language often expect words to be clearly separated in speech. When this expectation is not met, they may mis-segment the input, interpreting multiple words as a single unfamiliar item or failing to identify known words altogether (Brown and Hilferty, 2006). Research suggests that word boundary confusion is a primary cause of listening failure, as inaccurate segmentation prevents subsequent lexical and syntactic processing (Cutler, 2012).

2.4.2 Reduced Forms and Perceptual Difficulty

Reduced forms are a defining characteristic of connected speech and a significant source of listening difficulty for L2 learners. Function words and grammatical elements are frequently pronounced in weakened or contracted forms, resulting in pronunciations that differ substantially from those presented in instructional materials (Roach, 2009). As a result, learners may fail to recognize high-frequency words such as auxiliaries, prepositions, and negation markers during listening.

Studies indicate that learners often report "knowing the words" but being unable to hear them in spoken discourse (Field, 2008). This phenomenon reflects a mismatch between learners' stored lexical representations and the reduced phonological forms encountered in natural speech. Without explicit instruction in reduced forms, learners may interpret listening difficulties as a lack of proficiency rather than as a normal consequence of connected speech, which can negatively affect motivation and confidence (Cauldwell, 2013).

2.4.3 Fast Speech and Processing Constraints

Speed of delivery is another factor frequently cited by learners as a major obstacle to listening comprehension. Native or proficient speakers often speak at rates that exceed the processing capacity of L2 listeners, particularly when speech contains reduced and connected forms (Rost, 2011). However, research suggests that the perception of speech as "too fast" is often linked to phonological reduction rather than actual speech rate (Field, 2008).

Fast speech increases cognitive load by reducing the time available for perceptual decoding and integration of information in working memory. When learners are unable to process the speech signal efficiently, they may miss critical information and fail to construct a coherent mental representation of the message (Vandergrift and Goh, 2012). This difficulty is compounded in authentic listening situations, where learners have limited control over input and minimal opportunity for repetition.

2.4.4 Learner Perceptions of Spoken English

Learner perceptions play a significant role in shaping listening behavior and strategy use. Many L2 learners perceive spoken English, particularly native-speaker discourse, as unclear, overly rapid, or fundamentally different from the English they have learned in classrooms (Graham, 2006). These perceptions often stem from repeated experiences of listening failure caused by unfamiliar connected speech features.

When learners attribute comprehension difficulties solely to speed or accent, they may adopt ineffective coping strategies, such as focusing exclusively on key content words or abandoning efforts to decode the full message (Field, 2008). Over time, this can lead to listening anxiety and reduced engagement with spoken input. Research suggests that raising learners' awareness of connected speech processes can positively influence perceptions by helping learners reinterpret listening difficulties as manageable and predictable (Cauldwell, 2013).

2.4.5 Implications for Listening Instruction

The listening difficulties discussed above highlight the limitations of instructional approaches that emphasize comprehension outcomes without addressing the perceptual foundations of listening. Word boundary confusion, reduced forms, and fast speech are not random obstacles but systematic features of spoken English that can be addressed through targeted instruction. Explicit teaching of connected speech features has been shown to improve learners' decoding ability, increase confidence, and enhance overall listening comprehension (Field, 2008)(Celce-Murcia et al., 2010)

Understanding learner perceptions and the sources of listening difficulty is essential for designing effective pedagogical interventions. By focusing on connected speech as a core component of listening instruction, educators can better equip learners to process authentic spoken input and reduce the gap between classroom English and real-world communication.

2.5 Teaching Listening: Traditional and Explicit Phonological Instruction

Approaches to teaching listening in second language classrooms have evolved considerably over time; however, instructional practices often continue to prioritize comprehension outcomes over perceptual processes. Traditional listening pedagogy has largely focused on the development of strategies and increased exposure to spoken input, with relatively limited attention given to the phonological features that shape natural speech (Field, 2008). This section contrasts strategy-based instruction and exposure-only approaches with explicit phonological teaching, highlighting the strengths and limitations of each.

2.5.1 Strategy-Based Instruction

Strategy-based instruction aims to improve listening comprehension by equipping learners with cognitive and metacognitive strategies such as predicting content,

identifying key words, monitoring comprehension, and inferencing meaning from context (Vandergrift and Goh, 2012). These strategies are typically taught through pre-listening, while-listening, and post-listening activities, which encourage learners to engage actively with spoken input.

Research suggests that strategy training can enhance learners' awareness of listening processes and improve confidence, particularly for higher-level learners (Graham, 2006). However, critics argue that strategy-based instruction often assumes that learners can accurately perceive the speech signal in the first place. When bottom-up processing is compromised due to unfamiliar connected speech features, the effectiveness of top-down strategies is significantly reduced (Field, 2008). As a result, learners may rely on guessing rather than genuine comprehension, leading to partial or inaccurate understanding.

2.5.2 Exposure-Only Approaches

Exposure-only approaches to listening instruction are based on the assumption that repeated exposure to spoken language will naturally lead to improved comprehension over time. In such approaches, learners are encouraged to listen extensively to authentic materials such as lectures, conversations, or media, with minimal explicit instruction (Rost, 2011). This approach aligns with naturalistic theories of language acquisition, which emphasize input as the primary driver of learning.

While increased exposure is undoubtedly beneficial, research indicates that exposure alone is often insufficient for L2 learners, particularly in contexts where opportunities for meaningful interaction are limited (Field, 2008). Without guidance, learners may repeatedly fail to decode connected speech features, reinforcing inaccurate perceptual representations. Consequently, exposure-only approaches may lead to fossilization of listening difficulties rather than gradual improvement, especially for learners at lower proficiency levels.

2.5.3 Explicit Phonological Instruction

Explicit phonological instruction focuses on raising learners' awareness of the phonological characteristics of spoken language and providing guided practice in perceiving and processing these features. In the context of listening, this approach involves direct instruction in connected speech features such as assimilation, elision, contractions, weak forms, and flapping, combined with focused listening activities (Celce-Murcia et al., 2010).

Research increasingly supports the effectiveness of explicit phonological instruction for improving L2 listening comprehension. Studies have shown that when learners are taught how spoken forms differ from written forms, they are better able to segment the speech stream, recognize familiar words, and process input more efficiently (Field, 2008)(Cauldwell, 2013). Unlike strategy-based approaches, explicit phonological teaching strengthens bottom-up processing, thereby providing a more stable foundation for higher-level comprehension strategies.

Moreover, explicit instruction helps learners reconceptualize listening difficulties as systematic and predictable rather than as indicators of personal inadequacy. This shift in perception has been shown to reduce listening anxiety and increase learner engagement with authentic spoken input (Cauldwell, 2013).

2.5.4 Comparative Evaluation of Instructional Approaches

While strategy-based instruction and exposure-only approaches contribute to listening development, their effectiveness is limited when phonological processing is not adequately addressed. Explicit phonological instruction does not replace strategies or exposure; rather, it complements them by equipping learners with the perceptual tools necessary to decode spoken input accurately (Vandergrift and Goh, 2012).

In this regard, teaching connected speech features represents a pedagogically balanced approach that integrates bottom-up and top-down processing. By addressing the phonological foundations of listening, explicit instruction can enhance the effectiveness of other listening approaches and lead to more sustainable improvements

in comprehension. This perspective provides the pedagogical rationale for the present study, which investigates the impact of teaching connected speech aspects on L2 listening comprehension.

2.6 Empirical Studies on Teaching Connected Speech in the Pakistani ESL Context

Empirical research on listening instruction in Pakistan has traditionally prioritized reading, writing, and grammar, with listening often receiving minimal pedagogical attention. However, the research on the teaching of connected speech has grown steadily over the past two decades, reflecting increased recognition of the role of phonological processing in L2 listening comprehension. Within this limited body of listening research, studies focusing explicitly on connected speech instruction remain scarce. While earlier studies tended to focus on general listening strategies or pronunciation accuracy, more recent research has begun to investigate the effects of explicit instruction in connected speech features on learners' listening performance. This section reviews key empirical studies in this area, highlighting major findings, methodological approaches, and limitations. A growing number of Pakistani researchers have highlighted phonological and perceptual difficulties as major contributors to learners' poor listening comprehension, thereby indirectly supporting the need for explicit instruction in connected speech features.

2.6.1 General Listening Difficulties Identified in Pakistani ESL Research

Several Pakistani studies report that listening is consistently perceived by learners as the most difficult language skill. Mahmood (2018) found that university-level ESL learners struggled to comprehend spoken English despite adequate grammatical knowledge, attributing these difficulties to fast speech, unfamiliar pronunciation patterns, and reduced forms. Similarly, Rahman and Khan (2020) observed that

learners frequently failed to recognize familiar words in spoken texts, suggesting a perceptual rather than a lexical deficit.

These findings align with international research indicating that L2 listening problems often stem from weaknesses in bottom-up processing. A study conducted outside Pakistan emphasizes that "it is important that pronunciation instructors introduce phenomena related to phonological processes in English connected speech and encourage L2 learners to recognize and identify these features in native speech" (Rattansak, 2025). Rattansak further argues that awareness of such sound variations plays a crucial role in the comprehension of fast, natural speech. However, most Pakistani studies stop short of examining specific connected speech features, instead treating pronunciation and listening as separate instructional concerns.

2.6.2 Phonological Awareness and Listening Comprehension

A limited number of Pakistani studies have explored the relationship between phonological awareness and listening comprehension. Shahzad and Ali (2019) reported that learners with greater awareness of English sound patterns demonstrated better listening performance than those taught through traditional comprehension-based approaches. Their findings suggest that explicit attention to phonological features can enhance learners' ability to decode spoken input.

Similarly, Iqbal and Ahmed (2021) examined the effects of pronunciation instruction on listening comprehension and found modest but statistically significant improvements in learners' ability to process spoken English. Although their study did not focus exclusively on connected speech, it highlighted learners' difficulties with reduced forms and fast speech, indicating the relevance of connected speech phenomena in the Pakistani ESL context.

2.6.3 Limited Instruction on Connected Speech Features

Research suggests that explicit instruction in connected speech features such as assimilation, elision, contractions, and weak forms is largely absent from Pakistani

ESL classrooms. Hassan (2017) notes that pronunciation teaching in Pakistan is often restricted to isolated sounds and word-level accuracy, with little attention paid to how sounds change in continuous speech. As a result, learners remain ill-prepared to process authentic spoken English.

Khan and Mahmood (2020) further observe that listening activities in Pakistani classrooms are frequently assessment-oriented, focusing on answering comprehension questions rather than developing perceptual skills. This approach reinforces the misconception that listening difficulties arise solely from limited vocabulary or insufficient exposure, rather than from systematic phonological variation in spoken English.

2.6.4 Learner Perceptions and Classroom Realities

Pakistani learners consistently report that native or proficient speakers of English speak too fast or unclearly. While similar perceptions have been documented in global research, these difficulties are intensified in Pakistan due to limited exposure to authentic spoken input and minimal training in connected speech features. Learners often rely heavily on written English as their primary reference point, which further widens the gap between classroom learning and real-world listening demands.

Teachers, on the other hand, frequently report a lack of training and resources for teaching listening and pronunciation effectively (Hassan, 2017). This institutional constraint contributes to the continued marginalization of connected speech instruction within Pakistani ESL pedagogy.

2.6.5 Methodological Gaps in Pakistani Research

To address listening comprehension difficulties, Pakistani researchers have proposed various instructional strategies. Some studies emphasize the use of authentic materials in the classroom. For example, Shahid et al. (2021) argue that ESL learners' listening comprehension is hindered by limited schema development, as inadequate background knowledge restricts learners' ability to connect input with

meaning. In contrast, Waqas et al. (2025) advocate explicit instruction in weak forms and aspects of connected speech to improve pronunciation and listening. Pathan (2023), citing Goh, reports that (66%) of learners identified the speaker's accent as one of the most significant factors affecting comprehension.

Despite these insights, a major limitation of existing Pakistani research is the lack of experimental or quasi-experimental studies that specifically investigate the impact of teaching connected speech on listening comprehension and the reduction of difficulty related to connected speech perception and processing. Most studies rely on surveys, perception questionnaires, or general proficiency tests, offering limited evidence of causal relationships between instructional intervention and listening improvement (Mahmood, 2018)(Rahman and Khan, 2020).

Furthermore, few studies examine individual connected speech features or measure pre- and post-instruction listening gains. This methodological gap limits the ability to draw firm conclusions regarding the effectiveness of phonological interventions in the Pakistani ESL context.

2.6.6 Implications for the Present Study

The reviewed literature indicates that while Pakistani ESL learners experience significant listening difficulties related to phonological processing, explicit instruction in connected speech remains underexplored and underutilized. Existing studies identify the problem but rarely address it through systematic and targeted pedagogical intervention.

The present study seeks to address this gap by experimentally examining the effects of explicit instruction in connected speech features on listening comprehension among Pakistani ESL learners. By focusing on perceptual processes and employing a structured instructional approach, the study aims to contribute context-specific empirical evidence to an area that has received limited attention in Pakistani applied linguistics research.

2.7 Pakistani ESL Context and Listening Instruction

In Pakistan, listening is not systematically included as a skill in ESL instruction at most educational levels. While some elite private schools have recently begun to incorporate listening activities, the majority of university students have had little or no exposure to spoken English instruction during their earlier education. Given the limited instructional time allocated to English courses, extensive listening practice may not always be feasible. However, learners can be provided with targeted instruction that strengthens their linguistic awareness and enhances their ability to perceive spoken English.

It is often assumed that L2 learners will naturally acquire connected speech features as they gain fluency and proficiency, as their interlanguage develops (Rogerson-Revell, 2011). However, achieving this requires teacher training and pedagogical support. Waqas et al. (2025), in their study on teaching weak forms, highlight several challenges, including limited teacher expertise in phonetics and phonology, scarcity of instructional materials, and the need for ongoing professional development.

In such contexts, emerging technologies, including artificial intelligence, offer potential support. When integrated into the syllabus in a systematic and pedagogically informed manner, AI-mediated instruction may enhance learning outcomes by providing consistent input, adaptive feedback, and repeated practice opportunities. In settings where access to trained teachers is limited, such tools may partially compensate for gaps in instructional expertise.

Chapter 3

Research Methodology

3.1 Introduction

This chapter describes the study's methodological framework: the research design, participants, educational intervention, data collection methods, and analysis tools as well as theoretical and conceptual frameworks. The study assessed the perceptions of Pakistani ESL learners about connected speech and tested the efficacy of focused instructional interventions of weak forms, contractions, assimilation, elision, and alveolar flapping in improving listening comprehension. A quasi-experimental mixed-methods methodology was employed to assess both quantifiable alterations in listening performance and learners' perceptual reactions to the teaching of connected speech aspects.

3.2 Theoretical Frameworks

This was an experimental study that involved an integrated framework for developing listening comprehension through structured exposure and for reducing cognitive load. For this, two core frameworks had been selected based on their relevance to the teaching of connected speech and listening comprehension in ESL contexts.

3.2.1 Cognitive Load Theory

CLT was integrated into the study through the structured sequencing of lessons starting from less challenging level to most difficult. Understanding fluent speech places high processing demand on learners, therefore this graded exposure was intended to reduce cognitive overload and facilitate long-term retention. Listening input was increased in difficulty over twenty-four sessions, allowing learners to build familiarity with connected speech patterns without being overwhelmed. CLT justifies the spiral remedy by arguing that lesson plans should be such that reduces cognitive load and build schema gradually, which is exactly what my spiral model aimed to do. It supports my rationale for scaffolded listening training that starts with easier tasks and moves toward fluent speech decoding.

3.2.2 Relevance of Cognitive Load Theory to Connected Speech Perception Difficulty

Cognitive Load Theory (CLT), proposed by Sweller (1988), emphasizes that learning is optimized when instructional design aligns with the limitations of working memory. Working memory has a limited capacity to process new information, whereas long-term memory stores knowledge schemas that facilitate fluent comprehension. When information exceeds working memory capacity, cognitive overload occurs, hindering learning and comprehension.

In the context of connected speech processing difficulty, learners often experience intrinsic cognitive load due to the complex phonological modifications in connected speech such as assimilation, elision, contractions, weak forms, and flapping. These phenomena distort the expected lexical boundaries, increasing processing effort. Additionally, extraneous cognitive load may arise from poorly designed listening input, lack of metacognitive guidance, or unfamiliar accents.

This study drew on CLT to propose that systematic exposure to General American (GA) connected speech through controlled digital input, coupled with explicit awareness-raising activities which can reduce cognitive overload. By scaffolding

listening experiences and gradually increasing difficulty, learners' phonological schemas can adapt, shifting load from working to long-term memory.

3.2.3 Integration of Cognitive Load Theory into the Proposed Methodology

The qualitative, phenomenological design of this research aimed to trace participants' lived experiences of listening difficulty and interpret how cognitive overload is experienced during connected speech comprehension. Through in-depth interviews and reflective responses, the study sought to uncover the cognitive routes learners employ when decoding reduced speech forms. To operationalize CLT principles:

i. Intrinsic load: was moderated by selecting input that progresses from isolated features (e.g., elision) to natural speech sequences.

ii. Extraneous Load: was minimized by ensuring clarity of input, using familiar topics, and avoiding visual/linguistic distractions.

iii. Germane Load: (schema-building effort) was encouraged through reflective prompts and guided awareness tasks that help learners consciously reorganize their phonological perception.

Hence, the methodology was not only descriptive but strategically grounded in CLT, ensuring that participants' cognitive resources were directed toward meaningful learning rather than mere decoding effort (See figure 3.1).

3.2.4 Spiral Learning Approach

The instructional design of the present study is guided by the spiral learning approach, which emphasizes the systematic revisiting of key concepts over time while gradually increasing levels of complexity and cognitive demand. Rather than introducing connected speech features as isolated or one-time content items, spiral learning treats them as cumulative skills that are reinforced, extended, and integrated across successive instructional cycles.

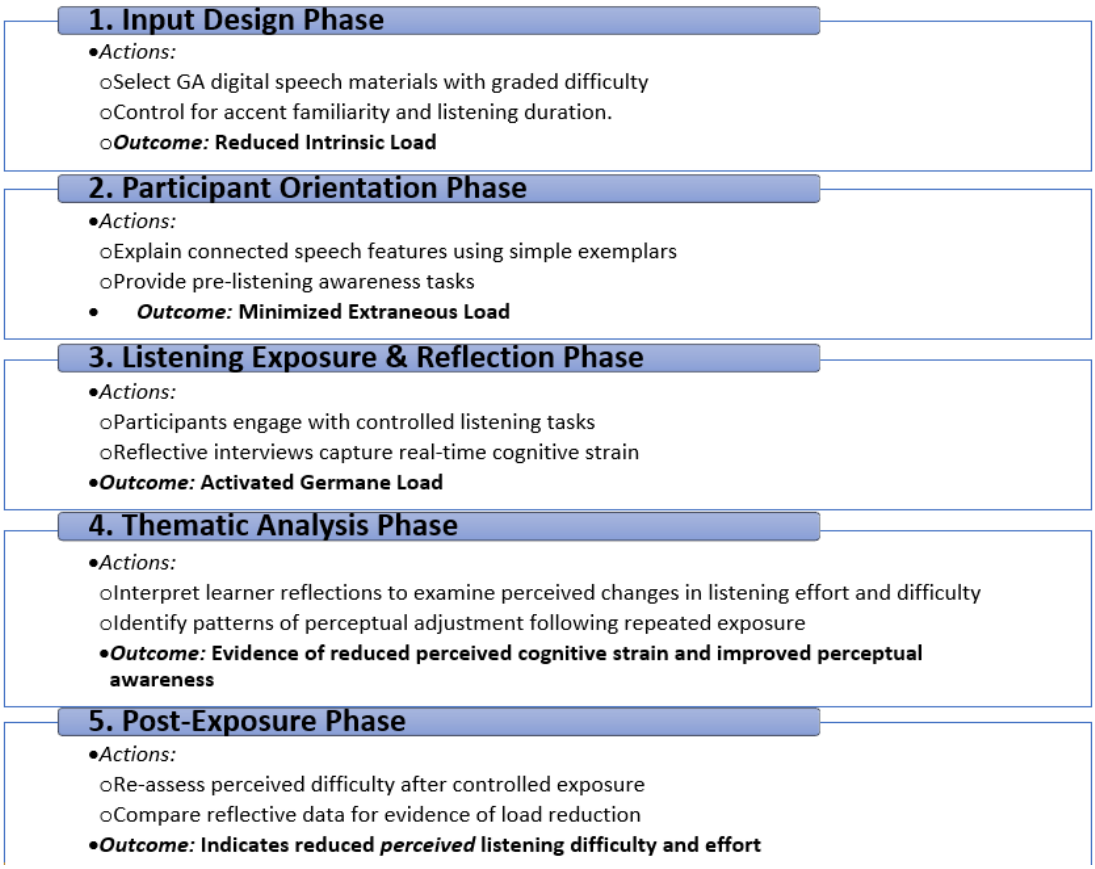


FIGURE 3.1: Flow Chart: Methodology Pathway to Reduce Cognitive Overload

Within this framework, learners are first exposed to connected speech features in simplified and highly controlled contexts. Initial lessons focus on recognition and awareness, allowing learners to familiarize themselves with reduced and modified forms of speech without excessive processing demands. As instruction progresses, these same features are reintroduced in new contexts that require higher levels of perceptual discrimination, contextual interpretation, and real-time processing. This cyclical return to previously introduced material enables learners to consolidate earlier learning while progressively developing more automatic listening skills.

Spiral learning also supports the integration of multiple connected speech features over time. Features such as weak forms, assimilation, and elision are not taught in isolation throughout the instructional period; instead, each feature is revisited and combined with previously learned features in later lessons. This cumulative progression reflects the natural characteristics of authentic spoken discourse, in which multiple phonological processes occur simultaneously. By encountering these

features repeatedly across varied tasks and listening texts, learners are better able to develop flexible and transferable listening strategies.

The lesson sequencing in this study reflects core spiral learning principles. Each instructional cycle follows a progression from guided perception to contextualized listening practice, with earlier lesson content systematically embedded into subsequent tasks. Previously taught features reappear in new listening passages, activities, and assessment tasks, ensuring continuity and reinforcement rather than linear completion. This design allows learners to deepen their understanding incrementally while reducing the likelihood of forgetting or fragmented knowledge.

Overall, the spiral learning approach provides a coherent pedagogical structure for connected speech instruction by aligning repetition with progression. Its emphasis on recycling, cumulative learning, and increasing complexity ensures that learners are not merely exposed to connected speech features but are given sustained opportunities to internalize and apply them in increasingly authentic listening contexts.

3.2.5 Supporting Pedagogical Principles

In addition to the main theoretical foundations, the instructional design of this study is informed by key pedagogical principles related to comprehensible input and learner noticing. Listening materials were selected and sequenced to ensure that spoken input remained accessible while gradually increasing in complexity, allowing learners to engage with connected speech features without becoming overwhelmed. At the same time, instructional tasks were designed to draw learners' attention to reduced and modified speech forms that are often missed in natural listening conditions. Through focused listening activities and guided practice, learners were encouraged to notice these features within meaningful contexts, supporting more effective perception and interpretation of authentic spoken discourse.

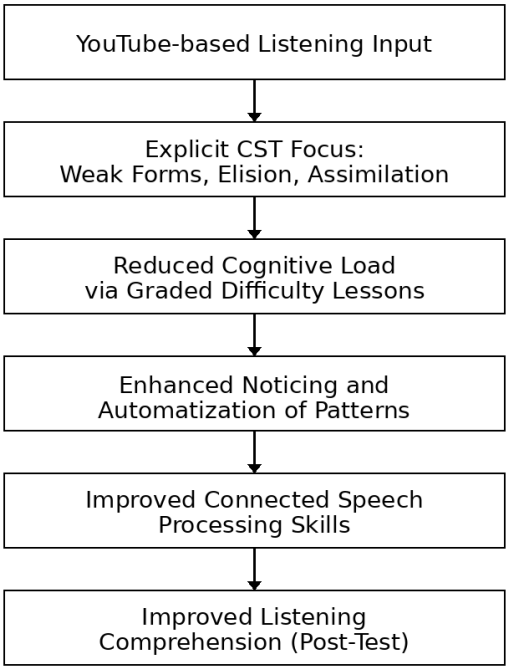


FIGURE 3.2: Visual Model of the Conceptual Framework

3.3 Conceptual Framework

The conceptual framework of this study integrates cognitive, perceptual, and instructional perspectives to address persistent listening difficulties faced by Pakistani ESL learners. These difficulties often stem from limited exposure to natural spoken English and an overreliance on written forms of the language. As a result, learners struggle to process fluent speech despite regular engagement with English-language media.

Grounded primarily in Cognitive Load Theory and models of listening comprehension, the framework emphasizes gradual skill development through structured exposure, explicit listening instruction, and repeated engagement with authentic input. Rather than treating listening difficulties as a result of insufficient exposure alone, the framework recognizes the need for guided perceptual training that helps learners adjust to the realities of connected speech (see figure 3.2).

The framework therefore serves both explanatory and pedagogical functions: it explains why connected speech poses challenges for ESL learners and outlines how these challenges can be addressed through systematic instructional design.

3.3.1 Justification for the Conceptual Framework

Listening difficulties in Pakistani ESL contexts are frequently linked to limited instructional focus on natural spoken English. While learners are often encouraged to consume English-language media, they are rarely taught how to listen effectively to such input resulting in comprehension difficulties. Consequently, many learners remain dependent on subtitles or become disinterested altogether.

Traditional textbook-based instruction typically prioritizes written grammar and vocabulary while neglecting the phonological characteristics of fluent speech. The proposed framework addresses this pedagogical gap by combining gradual exposure, authentic spoken input, and explicit listening instruction. The connected speech training module operationalizes this framework by sequencing instruction, managing cognitive load, and assessing perceptual development through pre- and post-intervention measures. Overall, the framework validates the module as a theory-driven instructional strategy rather than a purely intuitive teaching intervention. It offers a structured approach to addressing a documented listening comprehension gap in Pakistani tertiary-level ESL classrooms.

3.4 Research Design

This research employed a quasi-experimental mixed-methods methodology, incorporating quantitative pre-test and post-test data together with qualitative focus group input to investigate the learning of the ESL students through the designed intervention. The research was conducted as a classroom-based teaching sessions, which helped with the observation of learners' listening behaviors inside their natural learning setting for validity. This allowed the researcher to observe how ESL learners respond to the intervention and how much learning takes place in authentic teaching contexts.

The descriptive part of the study comprises of the systematic documentation of learners' responses to spoken input having features such as weak forms, contractions, assimilation, elision, and alveolar flapping. The study also classified student performance beyond basic distinctions of right and wrong, and included a range

of recognition, partial decoding, and misinterpretation with the help of carefully designed rubrics (See Appendix D). This method helped us comprehend how students process reduced speech. This study aimed to analyze changes in learners' perceptual performance post-intervention by identifying patterns of difficulty, approximation, and success in their post-listening responses.

An intervention-based methodology was incorporated into the research design to investigate whether explicit training may facilitate perceptual enhancement in learners' processing of related speech. Research on connected speech perception is still lacking in the setting of Pakistani ESL. Instead of just passively watching learners have trouble listening, the study specifically included the teaching of five very important connected speech aspects; weak forms, contractions, assimilation, elision, and alveolar flapping. The intervention proved useful in knowing how teaching connected speech can help ESL students understand what they hear.

3.4.1 Rationale for the Quasi-Experimental Design

This study used a quasi-experimental pre-test/post-test approach, applied in the natural classroom environment. The participants were selected from regular courses instead of forming a special group of specific kind of students. This helped compare and decide how well students listened before and after the intervention in natural learning environment. This method offered a pedagogically sound technique for assessing instructional impact in authentic ESL classrooms. Comparing pre-test and post-test assessments of control and experiment groups determined that variations in listening comprehension could be significantly credited to the intervention rather than mere incidental exposure.

3.5 Participants

The research was carried out at a private university in Pakistan, where English is taught as a second language and predominantly serves as the medium of instruction in academic settings. The participants were undergraduate students who were taking an English language course that was required for their degree. In this

educational institute, the main focus of the teachers in an ESL class is on reading, writing, and grammar, while listening and pronunciation get less attention in the classroom. Since most of these students came from government or non-elite private schools, and had mostly learnt English teacher-centered classroom settings rather than by talking or listening to native or near-native speakers for long periods of time, they mostly came across spoken English through digital media like films and TV shows on the internet. The students were only exposed to natural spoken English in limited settings; YouTube videos sent by a teacher which only a few could understand and most relied merely on subtitles to get the meaning. This background is important for the current study because students who do not have much experience with spontaneous spoken input often have trouble hearing connected speech features like weak forms, elision, assimilation, and flapping, which are not usually taught directly in the local curriculum.

3.5.1 Criteria for Participant Selection

The study's target group comprised second-semester undergraduate students aged 19 to 21 years, engaged in a mandatory English language course at the institution. The participants had previously attended government or non-elite private institutions, where English was predominantly taught as an academic discipline. Students were chosen based on whether they were enrolled in the course and whether they would be able to participate for the whole study. There were no exclusions of participation based on gender, academic performance, or linguistic background.

3.5.2 Group Assignment

The study included 32 undergraduate students in total. The participants were split into two groups: a control group with 16 people and an experimental group with 16 people. The experimental group received direct teaching in related speech features via a structured spiral-based listening module, whereas the control group continued with standard course instruction without specific listening intervention.

3.5.3 Sampling Method

The study utilized a purposive convenience sampling method. The participants were the researcher's students for the sake of accessibility and relevance to the research objectives. This sampling method was deemed suitable for a classroom-based quasi-experimental study, since it facilitated research inside an authentic and natural instructional environment which ensured practical viability.

3.5.4 Voluntary Participation and Informed Consent

Students might choose whether or not to take part in the study. Before data collection began, participants were told what the research was for, what the instructional activities would be like, and how the data would be collected. They were told that their grades or standing in the institution would not change whether they choose to participate or not, or leave at any time. All participants gave their written consent, and their answers were kept private throughout the study. (See Appendix J)

3.6 The Instructional Design

A variety of teaching techniques were employed to help students develop their ears and to give them regular exposure to related speech aspects. These methods transitioned from clear instructions, regular practice, and a slow transition to real-life listening situations. Explicit explanation and modelling were employed only after the first time they heard the audios. Instead of giving rules, short explanations were given to help students understand patterns they heard in the listening material, together with the teacher modelling the goal forms. This helped students connect altered or shortened acoustic forms with their standard lexical forms. To improve perceptual accuracy and make it easier to recognize related speech characteristics, controlled drilling and guided repetition were used in group and individually as well. After listening activities, students did brief, focused tasks that required them to repeat and tell the difference between sounds. This helped

them get acclimated to little phonological nuances without putting too much strain on their brains thus decreasing cognitive overload. These activities focused more on how well they listened than how well they spoke.

Songs were employed as an extra source of real input to give students more time to hear connected speech in a natural and interesting way. Because of its rhythmic structure and repeated patterns, songs let learners hear reduced forms, connecting, and assimilation many times in meaningful way, which helped them remember what they learnt (See Appendix C).

The lessons moved from controlled practice to contextualized listening tasks. At first, challenges were based on single phrases and brief sentences. Later, tasks challenged learners to understand related speech in dialogues. This progressive change allowed students to use perceptual abilities they had already learnt on real spoken information, which helped them go from concentrated recognition to real-world listening comprehension.

3.6.1 The Spiral Listening Module

The Spiral Listening Module is a tiered, developmental approach to teaching connected speech and reduced forms through incremental listening exposure. Instead of presenting complex spoken input at once, the model revisits key features across successive stages, allowing learners to process them with gradually increasing difficulty. This recursive structure is intended to support perceptual adjustment while keeping listening demands manageable (see figure 3.3).

The first spiral, “Awareness”, introduces learners to connected speech features through controlled input, with an emphasis on noticing rather than full comprehension. The “Practice” spiral provides repeated and structured listening tasks that reinforce perception of these features. In the “Application” spiral, learners are exposed to less controlled, more naturalistic listening input to encourage transfer beyond the classroom. Overall, the spiral progression aligns with principles of Cognitive Load Theory by increasing listening demands gradually, thereby supporting listening stamina and reducing persistent misperceptions of spoken forms (see figure 3.4).

FIGURE 3.3: Core components of the Spiral Method

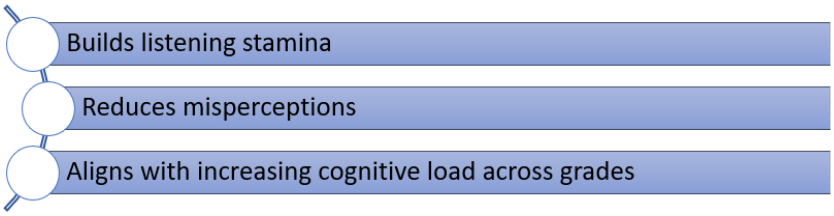


FIGURE 3.4: Benefits of the spiral methodology

3.6.2 Inductive Approach

An inductive approach was adopted in the instructional design of this study. Rather than presenting learners with explicit phonological rules at the outset, students were first exposed to examples of connected speech through guided listening tasks. Learners were encouraged to notice recurring patterns in reduced and linked speech, followed by structured practice and application activities. Explicit explanations were introduced only after initial exposure and practice. This approach aligned with the spiral listening module, as connected speech features were revisited across sessions with increasing complexity, allowing learners to gradually internalize patterns of natural spoken English without cognitive overload.

3.6.3 The shadowing technique

Originally developed in interpreting studies Lambert (1992) (Murphey, 2001), the shadowing technique has become a valuable strategy in second language teaching strategies for improving listening and speaking fluency. Shadowing involves learners modelling speech after hearing it which allows them to internalize rhythm, intonation, and the fluency of natural spoken language. Although shadowing is often associated with speaking practice, it was employed in the present study as a listening-focused technique rather than a speech production activity. The primary purpose of shadowing was to intensify learners' auditory attention to the speech stream and to support real-time processing of connected speech features such as reduction, elision, and weak forms. Research suggests that shadowing

enhances bottom-up processing skills by training learners' ears to detect reduced forms and linking in connected speech (Hamada, 2011). Moreover, it builds automaticity and gives confidence in oral production, as learners engage both receptive and productive skills in real time. In the present study, shadowing is not treated as a standalone approach but as a supportive technique embedded in the spiral-based listening module. Its purpose is to make the learners notice reduced forms and provide them with practical, repeated exposure to connected speech features in authentic contexts, thereby supporting the inductive learning process and reinforcing comprehension (Bruner, 1961).

A number of activities were developed to help the students notice, understand and practice the aspects of connected speech (see Appendix B). These were composed of brief listening clips, song-based audio, and organized worksheets that helped students notice, recognize, and practice in a controlled way. The main source of input was native spoken English, and the audios were chosen from YouTube having General American accent.

The listening samples were kept short so that students may hear them again and again and focus on certain phonological elements without getting overwhelmed by the influx of information. These excerpts were used in activities like guided listening, transcribing, and the comparison of spoken and written forms. The song-based materials were used to provide students more chances to hear related speech in a natural, repetitive and interesting way. This helped them learn about reduced forms, linking, and assimilation in longer conversations (see Appendix B).

Structured practice using worksheets and drill-based exercises helped students become more aware of their senses. This included tasks like filling in the blanks, short transcription, identification exercises, and drills. They were arranged according to spiral-based instructional design, which made sure that associated speech elements were reinforced in each session, and gradually increase in difficulty. Spoken output during shadowing was not evaluated for accuracy, nor was pronunciation instruction provided. Instead, shadowing functioned as a perceptual tool that supported the development of listening sensitivity and complemented the study's focus on connected speech perception.

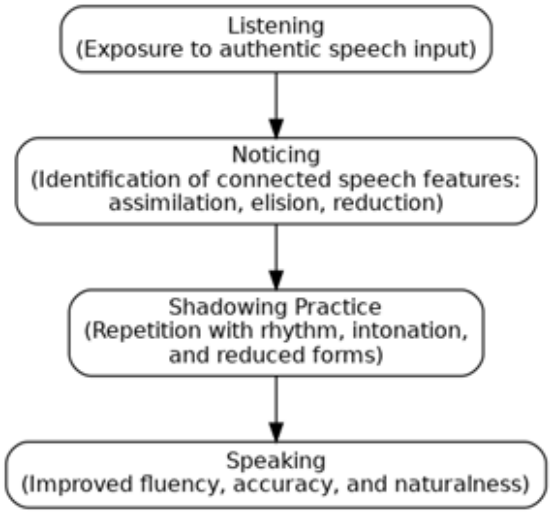


FIGURE 3.5: Flowchart of Shadowing Technique in Connected Speech Processing

3.6.4 Role of Noticing in Connected Speech Perception

The instructional design was also guided by the Noticing Hypothesis, which emphasizes the importance of learners becoming consciously aware of linguistic features in order for learning to occur. In the case of connected speech, features such as weak forms, elision, and assimilation often go unnoticed by L2 learners when they listen to fluent, real-time speech. To address this, the spiral listening module, inductive activities, and shadowing tasks were designed to gently draw learners' attention to these subtle features through repeated exposure and guided practice. Instead of relying on explicit rules, learners were encouraged to observe and recognize patterns in natural speech as they emerged across different listening tasks. This process of noticing helped support gradual development in learners' ability to perceive connected speech and worked alongside the instructional techniques used in the study.

3.7 Teaching Materials Selection Criteria

The choice of instructional materials was based on pedagogical and methodological factors that were important to the study's goals. First, clear audio input was prioritized to make sure that students could hear the associated speech qualities

they were supposed to while listening to real spoken English. To avoid cognitive overload, materials that were too quick or hard to hear were left out.

Second, how well it fit with the goal features was an important factor. We checked each listening extract and song to make sure that there were clear examples of weak forms, contractions, assimilation, elision, or alveolar flapping. This made sure that the teaching materials directly supported the goal of the intervention.

Finally, we chose materials that were both real and easy to understand. A slow to moderate pace of speech was recommended, since it allowed learners to practice natural pronunciation patterns while yet being easy for intermediate-level ESL learners to understand. Some of the listening items also contained relatively fast-paced speech. This was kept intentionally, as learners are likely to encounter similar speaking speeds in real-life situations outside the classroom. Including such input made the task slightly more challenging, but it helped ensure that the listening material reflected natural spoken English rather than carefully slowed or scripted speech.

The listening materials used in the intervention were mainly drawn from YouTube and were selected with considerable care. One of the primary concerns during selection was the suitability of the content for a Pakistani classroom context. For example, in one transcript a word was identified as culturally inappropriate, and it was therefore removed from both the audio recording and the written transcript before classroom use. Speech rate was another important factor in the selection process. To expose learners to different levels of listening difficulty, three audio texts were chosen to represent slow, normal, and fast-paced speech. The slow-paced extract came from the movie in which a child speaks to his mother, the second was a moderately paced dialogue from a popular movie, and the third was a fast-paced extract taken from backstage conversations among judges on a televised singing competition. Only the audio tracks were used during the intervention, as the videos were not shown to students in order to avoid reliance on visual cues. The selected audios were transcribed by the researcher and used as the basis for worksheet-based listening activities.

Songs were also incorporated into the instructional intervention as an additional listening resource to expose learners to naturally occurring connected speech in a familiar and engaging format. The selected songs were chosen with careful consideration of lyrical clarity, speech rate, and cultural appropriateness for a Pakistani classroom context. Preference was given to songs in which the lyrics were clearly articulated and where connected speech features such as weak forms, contractions, elision, and assimilation occurred naturally within the melody. Songs containing explicit language or culturally sensitive content were excluded. In order to vary listening difficulty, songs with relatively slower tempos as well as those with faster delivery were included. During classroom use, learners listened to the audio recordings without access to the music videos so that attention remained focused on auditory perception rather than visual support. Lyrics were transcribed by the researcher and adapted into worksheet-based activities, allowing learners to analyze connected speech features while engaging with authentic and motivating listening material.

3.7.1 Use of AI in Teaching Material

In addition to authentic audio extracts, targeted shadowing activities were incorporated into the intervention. For this purpose, the researcher developed short sentences aligned with the specific connected speech features being taught in each lesson. These sentences were then converted into audio recordings using text-to-speech technology in order to provide consistent and clear pronunciation models. The generated recordings approximated natural American English speech and were used to support learners' shadowing practice. During classroom implementation, students listened to the sentences and repeated them immediately after or simultaneously with the audio, depending on task difficulty. The use of these controlled audio samples allowed focused attention on individual phonological features while maintaining a natural-sounding speech model for practice.

3.8 Experimental Intervention

3.8.1 Overview of the Experimental Intervention

The instructional intervention took place over a period of sixteen weeks as part of the regular English language course. There were two lessons each week, each lasting about sixty minutes, for a total of thirty-two hours of instruction. The intervention took place in a regular classroom and was added to the existing course schedule so that normal lessons continued without disruption.

The intervention used a spiral-based teaching method, where associated speech elements were presented, reviewed, and reinforced throughout several sessions, with each one getting gradually more difficult. Each class had a set format that was based on CELTA-style lesson staging (see Appendix A). The first part of each session was a short question to get students thinking about what they already knew. Then, focused listening activities using short audio clips were introduced. After that, the learners did guided-noticing exercises, which drew their attention to certain elements of related speech in the input.

In the next steps, there were controlled practice activities like repetition, transcribing, filling in gaps, and short discriminating problems. These activities were meant to help students improve their perceptual accuracy and to get them to process spoken input from the ground up. Some sessions included shadowing activities that helped students learn how to imitate the natural rhythm and cadence of speech while also helping them remember reduced and connected forms.

During the intervention, the listening activities were carried out using audio materials that had been prepared in advance from selected online sources. Only the audio tracks were played in class, as the accompanying videos were deliberately excluded to ensure that learners relied solely on auditory input rather than visual cues. Each listening activity was accompanied by a worksheet designed around the audio text. Before listening, learners were briefly oriented to the task requirements. The audio recordings were then played multiple times, depending on the level of difficulty and learners' responses. In several instances, the recordings were paused

at specific points to give students sufficient time to write their answers and to reduce processing pressure. The pacing of the activities was adjusted according to the speech rate of the audio, with slower extracts requiring fewer repetitions and faster extracts requiring additional replay. Throughout the activities, learners focused on identifying and processing features of connected speech within the dialogues, while the teacher monitored comprehension and provided clarification when necessary.

During the intervention, the instructional input evolved from short phrases to full sentences. Previously introduced aspects were regularly repeated with fresh information to help reinforce the knowledge over time. Regular feedback was given to clear up confusions of the learners. One of the goals of the study was to help students become more aware of what to expect in real-life listening situations.

Within this overall instructional framework, specific connected speech features were systematically introduced and practiced through targeted listening, writing, and guided noticing exercises. The following subsection details how these activities were implemented during the treatment.

3.8.2 Treatment Procedures and Classroom Practices

The treatment was implemented as a learner-centered instructional intervention aimed at improving students' perception of connected speech in real-life listening contexts. Rather than treating connected speech as abstract phonological theory, the intervention focused on helping learners notice, hear, and gradually internalize how English sounds in natural, continuous speech.

All instructional sessions were planned following a CELTA-informed lesson staging framework, ensuring consistency and clarity across the intervention. Each lesson followed a structured sequence consisting of a lead-in, focused listening input, guided noticing, controlled practice, and feedback. This framework allowed learners to move systematically from awareness to practice while maintaining an appropriate cognitive load. While individual lesson plans were adapted according to the target connected speech feature and learner response, the overall structure remained consistent throughout the treatment. To avoid unnecessary repetition in the main

text, one comprehensive sample lesson plan illustrating the CELTA-based staging and activity sequence is provided in Appendix A.

Throughout the treatment, learners were repeatedly exposed to short stretches of spoken English and were guided to observe how words change, reduce, connect, or disappear in fluent speech. Each teaching session followed a consistent classroom routine. Lessons typically began with a brief lead-in activity designed to activate learners' prior knowledge and raise curiosity about spoken language patterns. This was followed by focused listening tasks using short audio clips presented without visual support.

Listening activities were accompanied by guided noticing exercises, where learners were encouraged to identify specific features of connected speech in the input. These were followed by controlled practice activities such as repetition, transcription, gap-filling, and listening discrimination tasks. In selected sessions, shadowing activities were incorporated to help learners follow the rhythm and pace of natural speech while reinforcing memory of reduced and connected forms.

All listening tasks were supported by worksheets specifically designed around the audio input. Audio recordings were played multiple times depending on the level of difficulty, and pauses were used strategically to reduce cognitive load and allow learners sufficient processing time. Instruction gradually progressed from short phrases to longer sentences, with previously taught features regularly recycled to support retention and cumulative learning.

3.8.3 Instruction of Individual Connected Speech Features

Each connected speech feature was introduced through a step-by-step instructional sequence that moved from awareness to controlled practice. Instruction typically began with simple examples and gradually progressed to more complex sentence-level input.

i. Weak Forms were introduced as a key source of listening difficulty for learners. Learners were guided to notice that common grammatical words are rarely pronounced in their strong forms in natural speech. Initial activities involved

identifying and underlining weak forms in short written sentences. These were followed by listening tasks in which learners marked reduced forms after hearing the audio. Writing exercises were used to visually reinforce vowel reduction and to help learners understand why certain words are often missed during listening.

ii. Contractions were taught as a natural feature of spoken grammar. Learners completed listening-based gap-fill exercises where contracted forms had to be reconstructed from audio input. Writing tasks required learners to transform full forms into contracted forms and vice versa, helping them develop form–meaning connections. Through repeated exposure, learners became more confident in recognizing contractions in fluent speech rather than relying on written forms.

iii. Assimilation was introduced through listening discrimination tasks that contrasted careful pronunciation with naturally assimilated forms. Learners were encouraged to recognize sound changes caused by neighboring sounds and to treat them as systematic patterns rather than pronunciation errors. Writing activities allowed learners to rewrite assimilated forms, reinforcing sound-based listening rather than spelling-based perception.

iv. Elision was taught as a feature where sounds disappear entirely in fast speech. Learners began with listen-and-repeat activities, followed by slow-versus-fast speech contrast drills. Writing exercises required learners to underline deleted sounds, select the most natural spoken form, and rewrite phrases as they are commonly heard in fluent speech. Reflection tasks encouraged learners to think about how elision affects listening comprehension.

v. Alveolar Flapping was introduced to familiarize learners with common American English pronunciation patterns. Through repetition drills and short listening tasks, learners learned to recognize flapped sounds in familiar words, reducing confusion during natural listening.

3.8.4 Combined Instruction of Connected Speech Features

In later stages of the intervention, selected connected speech features were taught in combination to reflect how they naturally co-occur in real spoken English.

During the second last teaching practice, weak forms and contractions were taught together due to their grammatical similarity and frequent co-occurrence. Learners worked with short dialogues containing multiple reductions within the same utterance. Writing tasks involved rewriting sentences into their natural spoken forms, while listening activities trained learners to identify multiple reduced forms simultaneously. This helped learners understand grammatical reduction as an integrated feature of fluent speech.

During the third last teaching practice, linking and assimilation were taught together to highlight sound interaction across word boundaries. Learners marked both linking points and assimilated sounds within the same sentences and verified their observations through listening. This integrated approach helped learners process continuous speech more holistically rather than treating each phenomenon in isolation.

Teaching combined features allowed learners to experience connected speech as a layered and dynamic system, closer to real-life listening conditions.

3.8.5 Pedagogical Orientation of the Treatment

The pedagogical orientation of the treatment was based on the assumption that listening difficulties in EFL contexts are often perceptual rather than linguistic. Learners frequently possess sufficient vocabulary and grammatical knowledge but struggle to recognize familiar language forms in fluent speech.

Accordingly, the intervention emphasized perception before production and focused on changing how learners listen rather than simply increasing exposure. Writing activities were integrated with listening tasks to support cognitive processing and help learners bridge the gap between spoken and written English.

The spiral nature of the treatment allowed learners to revisit previously introduced features in new contexts, reinforcing learning over time without overwhelming

them. By gradually shifting learners from word-by-word listening to flow-based listening, the treatment aimed to develop more effective and realistic listening strategies.

Overall, the intervention sought to move learners from text-dependent listening to speech-based listening, enabling them to engage with natural spoken English more confidently and efficiently.

3.9 Rationale for the Selected Connected Speech Features

The study mainly focused on five aspects of connected speech out of the many connected speech aspects that often make it challenging for ESL learners to decipher what they hear: contractions, weak forms, elision, assimilation, and alveolar flapping. These elements were chosen on the ground that they are common in natural spoken English and can make it hard to understand natural speech since they make word boundaries less clear. Besides, they are rarely taught in Pakistani ESL classrooms.

Short listening clips with several examples of the target phenomenon were used to introduce each element. The instruction followed a consistent order for all features. It started with real audio input and then moved on to guided noticing tasks that got students to pay attention to reduced or changed forms. Controlled recognition activities like transcribing, identification, and gap-filling exercises were done to improve perception, leading to ear development.

To help students recognize function words that are often unstressed in fluent speech, we talked about contractions and weak forms. Elision and assimilation were introduced to make learners more aware of how sounds can be deleted or changed across word borders. This brought to their awareness the intended lexical forms from changing acoustic information. Alveolar flapping was incorporated to help learners get used to the phonetic differences between American English's intervocalic /t/ and /d/ sounds, which sometimes leads to misidentifying common words. It is impertinent to mention here that alveolar flapping is also an aspect

of British English however less common than American English. Therefore, it was considered important from this angle that it will also help understand British English.

Throughout the sessions, previously taught aspects were addressed in combination with newly introduced material, facilitating the reinforcement of prior learning while accommodating the new information following spiral method. The major goal of the lessons was still to improve students' capacity to decode naturally occurring related speech while listening, not to improve their production accuracy.

3.9.1 Rationale for Including All Subtypes

All identified subtypes of connected speech features were included to ensure just coverage of the phonological phenomena of the study. Each subtype represents a distinct type of the same underlying process e.g., assimilation, elision, or reduction, and poses a different perceptual challenge for ESL learners (See table 3.1). Excluding particular subtypes would risk underrepresenting the variability of authentic speech, thereby limiting the ecological validity of the listening instruments.

Including all subtypes also supports content validity, as the pre-test and post-test systematically sampled the full range of realizations within each connected speech aspect rather than relying on isolated or prototypical forms. This approach ensures that learner performance reflects genuine listening competence rather than familiarity with a narrow set of examples.

From a pedagogical perspective, exposure to all subtypes aligns with a spiral learning approach, in which learners encounter the same phonological processes repeatedly across varied contexts and increasing levels of difficulty. Recycling subtypes in both tests allows for meaningful comparison of learner development while maintaining parallel test structure. Methodologically, this design strengthens the reliability and interpretability of pre–post gains by ensuring that improvements can be attributed to enhanced processing of connected speech as a whole, rather than to mastery of a limited subset of forms.

TABLE 3.1: Categories and Subtypes of Connected Speech Processes

Assimilation (8 subtypes)	**Alveolar Flapping (5 subtypes)**
Progressive assimilation	Intervocalic /t/ flapping
Regressive assimilation	Intervocalic /d/ flapping
Coalescent assimilation	Flapping across word boundaries
Place of articulation assimilation	Flapping before unstressed syllables
Manner of articulation assimilation	Flapping in multi-word phrases
Voicing assimilation	
Assimilation across word boundaries	
Complete vs. partial assimilation	
Weak Forms (7 subtypes)	**Elision (5 subtypes)**
Weak forms of articles (a, an, the)	Consonant elision (/t/ or /d/ deletion)
Weak forms of prepositions (to, for, of)	Vowel elision (schwa deletion)
Weak forms of auxiliary verbs	Elision in unstressed syllables
Weak forms of modal verbs	Elision across word boundaries
Weak forms of conjunctions (and, but)	Elision in rapid or casual speech
Weak forms of pronouns (him, her, them)	
Weak forms involving vowel reduction to schwa	

Contractions (5 subtypes)
Auxiliary verb contractions (I'm, you're)
Negative contractions (don't, can't,)
Perfect auxiliary contractions (I've, they've)
Modal verb contractions (he'll, we'd)
Tag-question and conversational contractions (aren't you?, isn't it?)

3.9.2 Rationale for Including Cross-Word and Within-Word Phonological Phenomena

In addition to cross-word connected speech phenomena, the instructional design also incorporated within-word phonological processes, particularly those occurring at morphemic and syllable boundaries. Words such as *glitter* and *input* exemplify how segmental changes, including elision, assimilation, and flapping, may occur within a single lexical item rather than across word boundaries. In glitter, for instance, the medial /t/ may be realized as a flap in natural speech, while in input, consonant clusters may undergo subtle articulatory adjustments that affect perceptual clarity.

Addressing these within-word processes is pedagogically significant because learners often expect words to be articulated in their citation forms and may fail to recognize familiar vocabulary when internal sound changes occur. Similarly, attention was given to phonological processes at word boundaries within phrases, where adjacent words interact to produce reduced or modified forms. Explicit inclusion of both within-word and inter-word phenomena aimed to develop learners' sensitivity to natural speech patterns, thereby enhancing listening comprehension and reducing cognitive load during real-time processing.

3.10 Data Collection Methods

The information for this study came from focus group discussion, as well as pre- and post-test evaluations. This strategy made it possible to carefully test participants' listening abilities both before and throughout the intervention.

The primary data sources included classroom observations, student replies and feedback, informal performance indicators for listening accuracy and engagement, and the researcher's personal notes. These tactics helped us study how learners used connected voice input and how the intervention made it easier for them to notice things.

3.10.1 Pre-Test and Post-Test Administration, Scoring, and Group Comparison

A pre-test and post-test design was used to evaluate the effectiveness of the instructional intervention for both the groups. Before the intervention, all participants took a listening pre-test to see how well they could understand and recognize features of connected speech in natural spoken English. The same format was used for the post-test at the end of the instruction period, but the listening material had slightly increased difficulty level. Overall, both the pre-test and post-test were designed to assess the same phonological categories; assimilation, elision, alveolar flapping, weak forms, and contractions using identical subtypes, thereby ensuring consistency in structure and scope (see Appendix C). Any variation in difficulty, therefore, stems not from test design but from the lexical choices and the degree of phonetic subtlety embedded in the examples. Across all categories, the pre-test generally employs more transparent and frequently encountered forms e.g., *dogs bark*, *next day*, *water, a cat*, *I'm happy*, whereas the post-test shifts toward more naturalistic, idiomatic, and compressed speech patterns such as *would you?*, *chocolate bar*, *water bottle*, and *shouldn't have gone*. These post-test items often involve less predictable collocations, multi-word constructions, and layered phonological processes that more closely mirror authentic spoken English. Consequently, the post-test can be characterized as moderately more difficult, as it demands a higher level of phonological awareness and listening sensitivity. This graduated increase in difficulty is pedagogically appropriate, as it allows the post-test to more effectively capture learners' progress following instruction and practice in connected speech.

The same rubric was used to score the learners' answers that looked at listening accuracy, recognition of reduced forms, and overall comprehension instead of just word identification. The experiment group's pre- and post-test performances were compared to identify changes resulting from the intervention, while the control group's scores established a baseline for natural progression without connected speech training. The analysis is organized around four comparative perspectives in order to provide a comprehensive evaluation of baseline equivalence, natural progression, instructional impact, and learning gains. This method made it possible

to find patterns of improvement, stability, or ongoing difficulty in both groups through comparison, which helped to support a structured evaluation of the instructional intervention.

3.10.2 Focus group discussion

Along with task-based responses, we also got feedback from students after lessons. Students were inquired about their thoughts on the difficulty of the aspect of study, and how much they thought their listening comprehension had improved at every stage. This input gave us a better idea of how learners felt about connected speech processing and added to the observational data by showing how learners saw their own progress and problems. This data was in form of recordings in Urdu language which were later translated and utilized in thematic analysis following (Braun and Clarke, 2006) Thematic Analysis (See Appendix F).

3.10.3 Classroom Observation

During the intervention, it was observed how engaged and responsive the participants they were when they did speech exercises. A key part of data collection was learner feedback which was comprised of written work from listening assignments, gap-filling exercises, and activities based on recognition that students did during lessons. It helped us identify patterns of correct recognition, partial decoding, and misunderstanding of associated speech elements. These responses included asking for clarification, correcting themselves, and interacting with others.

It was also noticed that students' confidence and concentration varied between sessions, particularly when they encountered previously learned material.

3.10.4 Triangulation of Data Sources

Classroom observations, learner feedback via focus group discussion, and pre- and post-test scores made triangulation possible. It helped study the patterns of perceptual change, rather than just one piece of evidence. The research amalgamated

various data sources to contextualize enhancements in listening comprehension within the broader scope of classroom interaction and pedagogical practices, providing a more comprehensive analysis of learners' advancement in connected speech processing.

3.11 Steps for Analyzing Data

The data underwent descriptive and thematic analytical techniques. The main goal of the study was to look for patterns in how students listened to connected speech and to see how their listening habits changed over time as they learned. The methodology focused on recording observed and reported changes in learners' recognition, segmentation, and interpretation of related speech elements, rather than merely producing numerical comparisons.

3.11.1 Pre-Test and Post-Test Analysis

This section outlines the procedures used to score, analyze, and interpret the learners' performance on the pre-test and post-test designed to assess connected speech processing difficulty. The analysis focuses on five core features of connected speech: assimilation, elision, flapping, weak forms, and contractions. Each feature was given equal weightage i.e, (20%) to maintain construct fairness and avoid disproportionate influence of any single category on the final proficiency score.

After the pre-test and post-test items were finalized, the audio recordings were produced. To ensure reliability and authenticity, the test scripts were recorded by a native speaker of American English. Each test item was played three times, with a three-second interval between repetitions, allowing participants sufficient time to process the input. When each phrase was played three times, the audio was briefly paused to allow participants to write down their responses. This procedure generated measurable data for subsequent analysis.

The data collected was analyzed through both quantitative and qualitative procedures to provide a comprehensive understanding of learners' progress. The pre-test

and post-test scores were compared using appropriate statistical measures to determine improvement in the recognition and comprehension of connected speech features, specifically contractions, flapping, assimilation, elision, and weak forms. The rubric allows for graded perceptual assessment rather than binary right–wrong scoring, capturing fine distinctions in auditory processing (See Appendix D).

3.11.2 Scoring Procedure

Learners' responses were scored using a five-point analytic rubric developed for this study. Each response was assigned a score between 1 and 5, where:

- 5 = Excellent.
- 4 = Good .
- 3 = Fair.
- 2 = Poor.
- 1 = Very Poor.

See Appendix D for the detailed rubric. The analytic approach was selected instead of binary scoring because it provides a more nuanced picture of learners' perceptual sensitivity to connected speech, especially in cases where students partially captured certain sounds but misperceived others.

3.11.2.1 Category-Level Scoring

Both pre-test and post-test data were entered into customized Excel scoring sheets. Item-level rubric scores were averaged within each connected-speech category, with all categories weighted equally. Each test item targeted one of the five connected-speech categories, and category means were calculated to represent learners' performance in that feature. Category scores were converted using the formula (Category Average / 5) $\times$ 20, ensuring equal contribution, which was (20%) from each category, and a maximum possible overall score of (100%).

3.11.3 Justification for Weighted Scoring Across the Aspects

All connected speech aspects investigated in this study; assimilation, elision, flapping, weak forms, and contractions were treated as conceptually equal components of connected speech perception, as no single feature was considered theoretically dominant. Prior research suggests that listening comprehension breakdowns typically result from the combined interaction of multiple connected speech features rather than reliance on a single element (Field, 2008) (Cauldwell, 2013). In line with assessment principles, rubric categories should only be differentially weighted when strong theoretical or empirical justification exists (Brookhart, 2013). Accordingly, no aspect was assigned greater conceptual importance than another.

However, at the level of score computation, weighted scoring was applied due to unequal numbers of test items across aspects. Specifically, the assimilation section comprised eight items, weak forms included seven items, while elision, flapping, and contractions each consisted of five items. Because of this imbalance, the use of a simple arithmetic average across aspect scores would have disproportionately inflated aspects with fewer items and understated those with more extensive item representation. To address this issue, aspect-wise average scores were converted into weighted scores proportional to their respective item counts prior to calculating the final listening score. This procedure ensured equitable contribution of each aspect to the overall score and preserved measurement validity.

It is pertinent to mention that although the test items were presented in short phrases rather than isolated words, marking focused only on the specific feature being tested in each item. The surrounding words were included simply to provide a natural listening context. For example, in *bread and butter*, the purpose was to check whether learners could notice the reduced form of *and*, not whether they could accurately identify every word in the phrase. Therefore, responses were scored based on recognition of the target feature, and minor inaccuracies in other parts of the phrase were not taken into account.

For data analysis, descriptive statistics were used to summarize pre-test and post-test performance and identify overall trends in listening development. Comparative

analysis examined aspect-wise gains following instruction or exposure. Improvement was determined through differences between pre-test and post-test mean scores, while individual score changes reflected learners' progress in specific connected speech features. In addition, qualitative analysis of learners' classroom responses and activities was conducted to identify recurring perceptual difficulties, learner strategies, and the effectiveness of the spiral-based instructional module in scaffolding listening competence.

Descriptive statistics were used for mean comparison, and inferential analysis involved paired-samples t-tests to examine within-group differences between pre-test and post-test scores.

3.11.3.1 Rubric for Pre and Post-test

The pre- and post-tests consisted of 30 listening items including all subtypes of the selected connected speech aspects (See table 3.2). The rubric ensured consistent evaluation across both tests and will help identify perceptual improvement in recognizing connected speech features. The pre-test and post-test were parallel in structure and content, as both targeted the same subtypes of connected speech features across all five aspects. Differences between the two instruments lay primarily in lexical choice and contextual complexity rather than in the phonological processes assessed, thereby ensuring comparability while allowing the post-test to reflect more natural and demanding spoken input. While the post-test included more complex lexical items and longer phrases to increase listening difficulty, full sentence contexts were avoided to reduce the possibility of guessing target features based on overall meaning rather than perceptual cues.

Table 3.2: Distribution of Items Across Connected Speech Categories

Category	**No. of Items**
Assimilation	8
Elision	5
Alveolar Flapping	5
Weak Forms	7
Contractions	5
Total	**30 Items**

Rubric for Scoring Recognition Accuracy. The five-point analytic rubric was designed to be applied to both the pre-test and post-test. This rubric was developed to capture not only the correctness of responses but also the degree of perceptual accuracy, allowing for a more nuanced assessment than a binary right–wrong marking scheme (See Appendix D). The rubric helped assess each of the thirty items in the pre-test and post-test. Each item score contributed to the learner's category average, which was then weighted;

$$\textbf{Final Score} = \sum \left(\frac{\textbf{Average Rubric Score per Category}}{5} \times \textbf{Category Weight} \right)$$

The rubric thus served a dual function: first, as a diagnostic tool in the pre-test to gauge baseline recognition, and second, as an evaluative measure in the post-test to determine perceptual improvement after exposure or training. The analytic design allowed partial credit for near-correct responses, ensuring that learners' gradual progress in connected speech perception could be quantitatively observed.

3.11.4 Final Score Calculation

The final score per participant was obtained using:

Final Score (100) = Assimilation (20%) + Elision (20%) + Flapping (20%) + Weak Forms (20%) + Contractions (20%)

Excel formulas were applied to automate the scoring, minimize human error, and allow clear visualization of learners' performance progress across tests.

3.11.5 Data Interpretation

Once all scores were generated, learners were assigned a Performance Level to categorize their proficiency:

- 90–100 = Excellent
- 75–89 = Good
- 60–74 = Moderate
- 40–59 = Developing

- 0–39 = Beginner

This classification enabled comparison of pre-test and post-test performance and provided insight into which connected-speech features learners struggled with the most.

3.11.6 Error Analysis

In addition to numerical scoring, learners' incorrect responses were reviewed to identify recurring patterns of misperception, such as vowel substitution, omission of weak forms, and confusion involving reduced consonants. These observations provided supplementary insight into how learners processed reduced and connected speech.

- Vowel-substitution errors
- Weak-form omission
- Flapping confusion
- Consonant-cluster reinterpretation
- Lexical substitution
- Full-segmentation breakdown

This analysis helped identify which features contributed most to connected speech difficulty and how learners processed reduced, assimilated, or weakened sounds.

3.11.7 Reliability and Validity

To maintain scoring reliability:

- All tests were scored twice by the researcher with a 48-hour gap.
- Discrepancies were re-evaluated to ensure consistent application of the rubric.

- The analytic rubric design aligns with the recommendations by Brookhart (2013) for performance assessment validity.
- Equal weighting ensures content and construct validity (Field, 2008; Cauldwell, 2018), preventing one feature from disproportionately influencing learners' overall scores.

3.11.8 Thematic Analysis and Pattern Recognition

After descriptive analysis, the qualitative data collected via focus group discussion given in Appendix F underwent thematic analysis. We listened to the audios and studied their comments over and over again to identify patterns in behaviors, emotions, and changes in perception. We came up with the first codes by looking at how they listened, like how they noticed reduced forms more, relied less on written guidance, and felt more confident throughout listening activities.

After that, these codes were put into larger analytical themes that showed how perceptions changed over time. We made sure that the themes were consistent and relevant by comparing them over and over again across sessions and data sources. This technique allowed the researcher to transition from discrete observations to more consistent interpretive categories that represented learners' developing engagement with connected speech.

3.11.9 Data Categorization for Analysis

The identified themes were systematically organized into coherent analytical categories that represented the key elements of listening development observed during the intervention. These categories included, but were not limited to, increased perceptual awareness of related speech elements, improved segmentation of spoken information, reduced misinterpretation of lexical boundaries, and greater learner engagement during listening activities.

Thematic organization of the data enabled a systematic examination of the progression of various aspects of listening comprehension over time and the influence of repeated instructional exposure on these advancements. This thematic approach

also made it easier to compare students and levels of training, which helped us make sense of qualitative data.

3.11.10 Connecting Analysis to Research Questions

The last step was connecting the themes to the research questions. We considered how patterns of learning related to the efficacy of the instructional intervention that helped students understand connected speech. It was ensured that the evidence from observations, learner responses, and pre- and post-test assessments was in sync with the study questions, and the data collection, analytical methods, and the study's goals were all in line with each other. This made the results more reliable and backed up fundamental conclusions about how focused connected speech teaching can help ESL students decode natural speech.

3.12 Ethical Considerations

This study was conducted with careful attention to ethical considerations in order to protect the rights, dignity, and well-being of all participants. Because the research took place in a regular classroom setting, participants were clearly informed about the purpose of the study, the nature of the instructional activities, and the methods of data collection. They were assured that participation was completely voluntary, their information would remain confidential, and they could withdraw from the study at any time without any academic consequences. Written informed consent was obtained before data collection began.

Participants' identities were protected throughout the study. Real names were not used at any stage; instead, codes were assigned, and all identifying information was removed from transcripts, observation notes, and written responses. The collected data were securely stored and accessed only by the researcher.

The study did not involve any form of marks for appearing in the tests. Participants were explicitly told that their performance in listening tasks or classroom activities related to the research would not affect their academic results. This approach

helped reduce pressure and anxiety and encouraged genuine participation in the learning activities.

Academic honesty was maintained throughout the research process. All teaching materials, audio resources, and theoretical frameworks were properly acknowledged, and the data were analyzed and reported accurately and transparently. The study followed institutional ethical guidelines and upheld principles of integrity and responsibility in educational research.

Chapter 4

Results and Discussion

4.1 Introduction

This chapter presents the results and discussion of the study. The findings are drawn from two sources of data: pre-test and post-test listening scores from the control and experimental groups are presented in Appendix C, and qualitative data obtained from focus group discussions with participants in the experimental group are presented in Appendix F. The purpose of this chapter is twofold: first, to examine the extent to which the proposed spiral-based teaching module enhanced learners' ability to decode fluent speech; and second, to explore whether this instructional approach reduced cognitive overload in ESL learners.

The chapter is divided into two main sections. Section 4.2 reports the quantitative results based on four comparative perspectives, while Section 4.3 presents the thematic analysis of learners' perceptions. Section 4.4 integrates and discusses both sets of findings in relation to the research questions.

The analysis is structured around four comparative perspectives:

1. within-group comparison for the control group
2. baseline equivalence between the groups
3. post-test comparison between groups
4. within-group comparison for the experimental group

This layered comparison allows for a comprehensive evaluation of natural progression, instructional impact, and qualitative changes in listening strategies. To ensure consistency and reliability of scoring, all responses were evaluated using a perception-based listening rubric (see Appendix E), which categorized responses into four levels: accurate recognition, minor error, phonologically plausible approximation, partial recognition, and misinterpretation. Each test item was rated on a scale from **1 to 5**, as shown in the rubric in Appendix E. This rubric helped assess each of the **thirty** items in the pre-test and post-test.

4.2 Quantitative Results: Pre-Test and Post-Test Comparison

4.2.1 Baseline Pre-Test Comparison:

Control Group vs Experimental Group. A comparison of the pre-test scores of the control and experimental groups demonstrates that both groups began the study at a comparable level of listening proficiency. As illustrated in Table 4.1, mean scores across all connected speech features showed only minor variation, with no systematic advantage observed for either group. Both groups displayed similar initial difficulties in perceiving native-like speech. Common patterns included frequent missegmentations, semantic substitutions, and incomplete comprehension across connected speech items. This baseline equivalence confirms that any subsequent change in post-test performance cannot be attributed to pre-existing differences in listening proficiency. Instead, observed differences must be interpreted in relation to the instructional conditions experienced by each group during the intervention period. Baseline Comparison of Control and Experimental Groups (Pre-Test Scores)

The comparison of pre-test results across the experimental and control groups revealed several common areas of difficulty, indicating that learners in both groups faced challenges in perceiving fluent spoken English. Analysis of aspect-wise performance showed particularly low scores in features associated with reduced

Table 4.1: Pre-test scores of both groups to establish baseline comparability prior to the intervention.

Feature	**Cntrl-Group Mean**	**Exp-Group Mean**
Contractions	3.19	2.64
Weak Forms	2.92	3.03
Elision	4.19	4.04
Assimilation	2.91	3.00
Alveolar Flapping	2.93	3.20
Overall Mean Score	**3.23**	**3.18**

and connected speech, such as weak forms, contractions, assimilation, and elision. These findings suggest that learners struggled not due to lack of vocabulary or grammatical knowledge, but due to difficulty processing phonological reduction and sound modification in continuous speech.

The pre-test results further indicated that learners relied heavily on orthographic expectations, which limited their ability to recognize familiar words when they appeared in reduced or altered forms. This mismatch between written knowledge and spoken input highlighted a gap in learners' listening competence and underscored the need for targeted perceptual training.

These identified problem areas directly informed the design of the instructional intervention. The treatment was therefore structured to focus on perception-oriented listening activities, guided noticing, and repeated exposure to connected speech features in a spiral manner. Specific classroom activities and exercises were selected to address the weaknesses observed in the pre-test, ensuring that the intervention was data-driven and pedagogically justified.

4.2.2 Post-Test Comparison: Control Group vs Experimental Group

The comparison of post-test results revealed a clear variation in learning outcomes between the experimental and control groups. While both groups demonstrated some level of improvement, the experimental group showed substantially higher gains across most connected speech aspects. In contrast, the control group's

improvement was limited and inconsistent, suggesting that exposure to regular instruction alone was insufficient to produce meaningful perceptual gains.

There was a considerable difference when the post-test scores of the control and experimental groups were compared. Table 4.2 reveals that the experimental group did better than the control group in all of the main related speech categories, namely elision, assimilation, weak forms, contractions, and alveolar flapping. The most obvious variations were in the parts of language that spiral education focused on, such as assimilation and elision. The control group still had some recognition and a lot of misinterpretation, but the experimental group was better at recognizing things, had greater phonological plausibility, and decoded things more consistently across items. After the test, the two groups did considerably differently, which shows that their listening outcomes were very different after the instructional session. It also shows that the disparities we noticed are related to the unique teaching method used.

Aspect-wise analysis of the post-test results indicated that learners in the experimental group performed notably better in features directly targeted during the intervention, particularly those involving phonological reduction and sound interaction.

Table 4.2: Post-test Listening Performance of Both Groups Following the Instructional Period

Feature	**Cntrl-Group Mean**	**Exp-Group Mean**
Contractions	2.83	3.98
Weak Forms	3.46	4.19
Elision	3.44	4.36
Assimilation	2.94	3.76
Alveolar Flapping	2.71	4.28
Overall Mean Score	**3.08**	**4.11**

These gains suggest that focused instruction and repeated exposure played a significant role in improving learners' ability to process fluent spoken input. The control group, which did not receive targeted connected speech training, continued to show difficulty in recognizing reduced and connected forms despite completing the same assessment. Overall, the post-test comparison reflects variation in learning

attributable to the instructional treatment, indicating that systematic connected speech instruction had a measurable impact on learners' listening performance.

4.2.3 Experimental Group: Pre-Test vs Post-Test Comparison

The results of the experimental group's pre-test and post-test show that the participants' perception of connected speech increased much after the intervention. The experimental group's pre-test results show that they know a little bit about how speech qualities are connected to each other. They did better at flapping and elision. But the lower mean scores in contractions demonstrate that the students weren't particularly good at using less prevalent grammatical forms before they obtained help with their grammar. Table 4.4 shows that students improved at recognizing reduced speech patterns. The most significant enhancements were noted in phonologically intricate elements such as assimilation, elision, and alveolar flapping, which had previously posed the greatest levels of confusion.

Individual gain scores revealed that *EG10* and *EG12* demonstrated the greatest improvement from pre-test to post-test, with gains of *45.01* and *32.51* points respectively, indicating substantial enhancement in connected speech perception following instruction as seen in Table 4.3.

Table 4.3: Gain scores of two participants with the greatest improvement

Group	**Pre-test Score**	**Post-test Score**	**Gain Score**
EG10	36.59	81.60	+45.01
EG12	51.69	84.20	+32.51

The responses after the test were more consistent, had less extreme distortions, and had a higher percentage of phonologically plausible approximations. They also improved at identifying words. These patterns, as illustrated in table 4.4, show that students paid more attention at one assessment point than at the other. They also suggest that the changes are more than just improved scores; they show that participants processed relevant speech in a more steady and structured way.

The average score of 4.11 shows that the educational intervention improved the students' understanding of the connected speech.

TABLE 4.4: A descriptive comparison of the experiment group's listening performance before and after the intervention

Feature	**Pre-test Mean**	**Post-test Mean**
Contractions	2.64	3.98
Weak Forms	3.03	4.19
Elision	4.04	4.36
Assimilation	3.00	3.76
Alveolar Flapping	3.20	4.28
Overall Mean Score	**3.18**	**4.11**

A comparison of the pre-test and post-test results of the experimental group revealed a marked improvement in learners' perception of connected speech following the instructional intervention (see Appendix C7). The post-test scores were consistently higher than the pre-test scores across all assessed aspects, indicating that learners benefited from the targeted instruction. This within-group comparison suggests that the instructional design was effective in addressing the listening difficulties identified at the pre-test stage.

Aspect-wise analysis further demonstrated that the most substantial gains occurred in connected speech features that were explicitly and systematically taught during the intervention. Learners showed notable improvement in recognizing weak forms, contractions, and features involving sound interaction such as linking and assimilation. Improvements were also observed in elision-related items, which had previously posed considerable difficulty at the pre-test stage. These results indicate that repeated exposure, guided noticing, and controlled practice contributed to enhanced perceptual accuracy. By gradually increasing task complexity and recycling previously introduced features, the intervention enabled learners to develop more realistic listening strategies and reduce reliance on orthographic expectations.

Secondly, the intervention was more about listening to understand than about speaking. This allowed students to focus on decoding speech as it flows spontaneously, without worrying about how they sound like a native speaker. Shadowing,

guided observation, and comparing transcripts are some of the methods that helped students link reduced spoken forms to their underlying lexical representations.

Third, students were more likely to employ meaning reconstruction instead of decoding words one at a time when they listened to real-life examples and conducted task-based listening tasks. This points towards how the ESL learners listen is especially important for Pakistani ESL students, who are often proficient in reading and writing but have trouble with spoken information in real time. This steadily changed their perception, and is in accordance with emerging theories about how the brain works that claim learners need to be slowly exposed to new information in order to keep their processing load under control.

4.2.4 Control Group: Pre-Test vs Post-Test Comparison

Comparing the control group's pre-test and post-test scores revealed only borderline improvement in overall listening performance. As shown in Table 4.5, slight improvement was observed in certain connected speech features, particularly contractions and weak forms; however, these improvements were inconsistent. In contrast, features involving phonological modification, most notably assimilation, elision, and alveolar flapping remained challenging for most of participants. High rates of misinterpretation and low levels of phonologically sound responses were detected, indicating minimal perceptual adjustment over the testing period.

Table 4.5: A descriptive comparison of the control group's listening performance before and after the instructional period

Feature	**Pre-test Mean**	**Post-test Mean**
Contractions	3.19	2.83
Weak Forms	2.92	3.46
Elision	4.19	3.44
Assimilation	2.91	2.94
Alveolar Flapping	2.93	2.71
Overall Mean Score	**3.23**	**3.08**

This result suggests that general exposure to English, without explicit instructional focus on connected speech, was insufficient to bring about meaningful changes

in learners' processing of natural spoken input. The limited progress in the control group indicates that connected speech features are unlikely to be learnt through routine English instruction at tertiary-level. These findings prove Field's (2008) assertion that the perceptual demands of connected speech require targeted pedagogical intervention rather than reliance on exposure alone.

4.2.5 Features Contributing to Connected Speech Perception Difficulty

A review of learners' responses across the pre-test and post-test showed that some connected speech features were noticeably harder to perceive than others. The results suggest that difficulties were not evenly distributed across features, with certain types of reduction causing more consistent problems for learners. The greatest difficulty was observed in items involving weak forms of function words. When words such as *and*, *to*, *can*, and *have* were reduced and unstressed within phrases, learners often failed to notice them or treated them as part of neighboring words. In many cases, these items were missed altogether, indicating that reduced stress and vowel quality made them particularly difficult to perceive in continuous speech.

Items involving reduced realizations of /t/ also posed considerable difficulty. Learners frequently struggled with cases where /t/ was realized in a weakened form in natural speech, such as in intervocalic contexts or across word boundaries. These reduced forms appeared to interfere with learners' expectations of clear segmental cues, making word recognition less straightforward.

Assimilation contributed to perception difficulty mainly when sound changes occurred across word boundaries. While learners sometimes misperceived these items, they were often still able to recover the intended meaning from the surrounding context. As a result, assimilation-related difficulties tended to be more limited in scope compared to those caused by more heavily reduced features.

Elision was also associated with some degree of perceptual difficulty, particularly in consonant clusters where one or more sounds were absent. However, the impact of

elision was generally more localized, affecting individual words rather than leading to broader breakdowns in phrase-level understanding.

Overall, the results indicate that features involving greater reduction and loss of prominence, especially weak forms and reduced realizations of /t/, were more likely to interfere with learners' perception of connected speech than features such as assimilation and elision, which tended to preserve clearer phonetic cues.

4.2.6 Variation in Learning Attributable to Instructional Strategies

The results show that teaching connected speech in an organized fashion can help students move from the English they learn in class to the English they speak in real life. The experimental group's progress indicates how vital it is to teach related speech in ESL listening programs for college students. It provides evidence of the effectiveness of the instructional design in improving learners' ability to process fluent spoken English. While individual variation in learning was observed, the general upward trend in scores suggests that structured connected speech instruction can play a meaningful role in enhancing listening comprehension.

Quantitative analysis of pre-test and post-test scores revealed clear variation in learning outcomes between the experimental and control groups, attributable to differences in instructional strategies. The experimental group demonstrated substantially higher mean gains across connected speech aspects, with consistently larger pre-to-post score differences than those observed in the control group. In contrast, the control group exhibited only modest mean score increases, reflecting regular progress rather than targeted perceptual development.

Aspect-wise comparison further indicated that the largest gains in the experimental group occurred in features that were explicitly addressed through structured listening, guided noticing, and repeated exposure, whereas gains in the control group were smaller and less consistent across aspects. The magnitude and consistency of score improvement in the experimental group, relative to the control group, quantitatively reflect variation in learning associated with the instructional strategies employed during the intervention.

4.3 Qualitative Findings: Thematic Analysis of Focus Group Discussions

4.3.1 Introduction to Thematic Analysis

The analysis aimed to explore learners' perceptions of the listening module and its impact on their listening and speaking skills. Focus group discussions were audio-recorded with participants' consent and transcribed verbatim in Urdu. The transcripts were then translated into English for analysis The focus group data were analyzed using a thematic analysis approach, following the principles outlined by Braun and Clarke (2006) using ATLAS.ti. To retain participant anonymity, all identifying information was removed, and participants were assigned codes (P1, P2 and so on). Due to the transcription being lengthy, only translated excerpts relevant to the identified themes are reported in the findings.

4.3.2 Theme 1: Perceived Listening Difficulties Prior to Instruction

Participants consistently indicated that prior to the educational intervention, they found it challenging to comprehend rapid, colloquial English speech, particularly in the context of international news, films, and dialogues featuring non-native speakers. Learners often said they felt like they couldn't understand, especially because of contractions and peculiar ways of saying things.

"When we used to listen to news from other countries, there were contractions we didn't understand." (P1)

"Before these sessions, we couldn't understand what foreigners were saying or films in English." (P8)

This theme shows that students didn't hear related speech very often and relied on meticulous, textbook-style pronunciation before the intervention.

4.3.3 Theme 2: Increased Awareness of Connected Speech Features

A prominent theme in most of the responses was the development of awareness regarding connected speech features such as contractions, reduced forms, and accent variation. Learners reported that aspects which previously caused confusion became clearer after systematic practice.

"After this, contractions and related features became clear." (P8)

"We learned many things that we had never been taught before." (P3) This suggests that absolute focus on connected speech helped learners decipher fast speech as systematic rather than random or careless.

4.3.4 Theme 3: Improvement in Listening Comprehension

The majority of participants in the module reported enhancements in their listening skills. Students claimed they understood spoken English better when it was in media settings and when they were talking to someone in real time. This shows that they went from not being sure how to decode natural speech to being more confident of how to do so.

"Our listening improved a lot" (P1)

"We can now understand a lot more when we watch films." (P5)

This stated improvement aligns with the measurable post-test advancements observed in the experimental group.

4.3.5 Theme 4: Transfer of Listening Gains to Speaking Skills

Although the main focus of this study was listening comprehension rather than speech production, the qualitative findings suggest that the intervention also had a noticeable impact on how learners felt about their own spoken English. Thematic analysis of classroom discussions and learner reflections showed that

many participants believed their speech had become more fluent and natural over time. Several learners reported greater awareness of how words connect in natural speech. Even though explicit speaking practice was not a core part of the instructional design, increased exposure to connected speech appeared to influence how learners perceived and monitored their own speech. These observations point to a possible link between improved listening perception and spoken fluency. However, since speech production was not formally assessed in this study, these improvements are reported as learner perceptions rather than measured outcomes.

"We can hear better and give the right answer." (P2)

"It helped me talk and listen better." (P16)

This demonstrates the connection between listening and speaking skills, particularly during interpersonal communication.

4.3.6 Theme 5: Positive Learning Environment and Pedagogical Value

Students expressed that a supportive classroom was very important and that they really liked the way the teacher taught. Participants also emphasized the need to integrate such modules into regular English instruction, preferably from an early educational stage.

"The class environment was friendly, which made it easy to understand." (P7)

"This type of module should be added to regular English classes." (P14)

This subject emphasizes the perceived educational significance of connected speech instruction.

4.3.7 Summary of Qualitative Findings

The qualitative findings substantiate the quantitative results of the study by illustrating that learners not only enhanced their test performance but also had

significant transformations in listening awareness, strategy utilization, and communication confidence. The alignment of test scores and learner views enhances the validity of the instructional intervention.

4.4 Integrated Discussion of Findings

4.4.1 Impact of Explicit Connected Speech Instruction

Across both the control and experimental groups, the pre-test data revealed considerable difficulty in decoding naturally occurring connected speech. Learners in the control group frequently misinterpreted target phrases as unrelated lexical items. This was especially clear when shortened versions were confused with phrases that had nothing to do with them for example, "dogs biting" was thought to mean "horse riding" or "stop fighting", which showed that bottom-up phonological processing was not working.

The experimental group's post-test findings, on the other hand, revealed big improvements in how well they recognized sounds, how well they partially decoded things, and how plausible their phonological responses were. Learners did not always get everything right, but they did get better at making guesses that were close to the right syllabic structure, stress patterns, or phonological aspects of the target phrases. This indicates that the intervention promoted perceptual reconfiguration instead of simple memorization.

This change backs up the idea that listening comprehension is not something that happens on its own, but rather something that can be taught through rising awareness and guided exposure. In Pakistan, it is very important for ESL learners to learn how to understand English, hence it is proposed that ear development be a part of the ESL curriculum.

4.4.2 Contextual Implications for Pakistani ESL Learners

The findings indicate that the spiral-based instructional approach enhanced learners' ability to decode fluent spoken English while also creating a low-pressure learning

environment. By reducing complexity through gradual sequencing, systematic recycling of features, and an emphasis on noticing and practice rather than rule memorization, the intervention minimized cognitive overload and led to measurable improvements in listening performance. These results provide strong pedagogical support for integrating connected speech instruction into ESL curricula, particularly at the university level. Based on the findings, teachers are encouraged to:

- provide explicit guidance on reduced and contracted forms in spoken English;
- recycle connected speech features through spiral sequencing;
- incorporate perception-based listening activities that prioritize phonological discrimination;
- emphasize learner noticing of patterns in continuous speech; and
- delay production-focused demands until learners demonstrate sufficient perceptual control.

The effectiveness of this approach across proficiency levels suggests that spiral-based connected speech instruction is broadly applicable. It is especially well suited to the Pakistani ESL context, where learners often demonstrate strong reading skills but experience persistent difficulty processing spoken English in real time.

A key implication of the study lies in clarifying a distinction that is frequently blurred in local ESL research. Much existing work assumes that listening competence develops naturally through exposure and extensive listening practice. While theoretically sound, this assumption presupposes sustained learner motivation and access to input conditions that are not always realistic in formal ESL contexts. As a result, listening is often treated as a skill that improves with increased input rather than as one that depends on specific phonological processes involved in real-time speech decoding.

The present study adopts a different perspective by framing listening difficulty in terms of connected speech processing difficulty. From this viewpoint, comprehension problems often arise not from limited vocabulary or grammar, but from difficulty perceiving and segmenting continuous speech. Even linguistically competent

learners may struggle when reduced or modified forms cannot be recognized in real time. By targeting connected speech processing difficulty directly, the study addresses an aspect of listening that has received limited attention in Pakistani ESL research.

Importantly, embedding listening and connected speech instruction in particular within the formal curriculum offers clear instructional benefits. When listening is part of assessed coursework, learners receive structured guidance and sustained practice rather than relying on informal exposure alone. Such integration promotes equitable access to listening development and supports a more inclusive and pedagogically grounded approach to ESL instruction.

4.4.3 Assimilation: Transitioning from Semantic Inference to Phonological Awareness

Prior to instruction, learners frequently interpreted assimilated forms as unrelated lexical items. Target phrases such as "dogs chase," "this year," and "green beans" were often misheard as semantically distant expressions, indicating that learners did not perceive assimilation as a systematic phonological process. Instead, reduced forms were treated as unfamiliar lexical input.

Following instruction, a clear shift in perceptual behavior was observed. Learners demonstrated improved segmentation by correctly identifying word boundaries, underlying phoneme sequences, and segmental substitutions in assimilated forms. In the post-training worksheet, most learners achieved near-perfect performance across different types of assimilation, with remaining errors reflecting phonologically plausible approximations rather than unrelated guesses.

This pattern reflects a transition from reliance on semantic inference toward phonological approximation, suggesting developing sensitivity to connected speech processes. Consistent with previous research, explicit instruction in the relationship between phonological processes and surface speech forms appears to support learners' ability to map reduced auditory input onto stable lexical representations. Importantly, systematic but imperfect approximations indicate perceptual development rather than failure, as connected speech perception improves gradually.

4.4.4 Elision: Improved Inference of Reduced or Missing Segments

Elision posed a significant challenge in the pre-test, with learners frequently failing to detect missing segments and instead interpreting elided forms as unrelated lexical items. For example, *hit 'em all* was mostly misperceived as semantically distant expressions, suggesting reliance on surface phonetic similarity rather than reconstruction of underlying forms.

Post-instruction analysis revealed a shift in learners' perceptual strategies. While learners consistently identified stressed lexical items, variability remained in the perception of reduced function words. Importantly, incorrect responses were phonologically plausible rather than random, indicating increased sensitivity to reduction, elision, and consonant cluster simplification (see Appendix H).

Errors were concentrated at points where word boundaries were obscured by connected speech, suggesting active attempts to segment continuous input. Compared with pre-training responses, post-training approximations remained closer to the target phonological structure, reflecting a move away from semantic inference toward bottom-up decoding. Consistent with models of L2 speech perception, such constrained approximations signal emerging perceptual competence rather than breakdown.

4.4.5 Alveolar Flapping: Reduction in Lexical Confusion

At the pre-test stage, alveolar flapping proved particularly challenging. Learners often misinterpreted flapped realizations of intervocalic /t/ and /d/, leading to substantial lexical confusion. These patterns suggest that learners initially treated flapping as unpredictable variation rather than a rule-governed phonological process.

Following instruction, learners demonstrated increased awareness of flapping environments and greater sensitivity to phonological constraints. Although some inaccuracies persisted, responses were more systematically constrained, and semantically unrelated interpretations were markedly reduced. This indicates growing

acceptance of phonetic variability within the sound system and recognition that a single lexical item may surface in multiple forms.

4.4.5.1 Emergent Pedagogical Insight: L1-Based Scaffolding and Phonological Overgeneralization

An important pedagogical insight emerged during instruction. The flapped realization of intervocalic /t/ and /d/ closely resembles the Urdu alveolar tap *ray*, a familiar sound for learners. Drawing attention to this similarity initially facilitated perception and production by reducing cognitive and articulatory load.

However, some learners overgeneralized this correspondence, temporarily replacing both flapped forms and the English rhotic /r/ with a single L1-based sound. This reflects a common interlanguage strategy in which learners simplify unfamiliar contrasts by collapsing them into familiar categories.

To address this, targeted contrastive practice was introduced, juxtaposing flapping contexts with words containing /r/, supported by explicit explanation and guided feedback. Over time, learners refined this distinction, reducing inappropriate transfer. Pedagogically, this highlights the value of L1-based scaffolding when used judiciously and paired with explicit awareness-raising to prevent phonological overgeneralization.

4.4.6 Weak Forms: Sensitivity to Function Word Reduction

Weak forms proved to be one of the most persistent sources of difficulty for learners, particularly when reduced grammatical items occurred in unstressed positions and carried low lexical salience. In the pre-test, learners frequently misinterpreted function words, often reconstructing them as semantically implausible lexical items. For instance, reduced sequences such as *I would've been* were sometimes perceived as unrelated strings like *out of bin*, suggesting that learners were attempting to match the acoustic input to familiar words rather than recognizing weak forms within a grammatical sequence.

During the intervention, learners were explicitly introduced to the schwa phoneme and its IPA representation /ə/ to raise awareness of its role as the default vowel in unstressed syllables. Instruction focused on high-frequency grammatical words such as *from*, *for*, and *to*, which learners often expect to hear in their strong forms. By repeatedly linking these items to schwa across different phonological environments, learners were encouraged to perceive vowel reduction as a regular and systematic feature of connected speech.

Following instruction, learners demonstrated greater sensitivity to weak forms, particularly in recurring patterns such as *a little noise*, "from the start," and *there are some apples* . Post-intervention worksheets showed that learners were increasingly able to preserve the intended syntactic structure of utterances, even when phonetic detail was incomplete or orthographically inaccurate. Instead of producing unrelated guesses, learners' responses reflected growing awareness of how reduced grammatical elements function within continuous speech.

These findings indicate that explicit instruction in weak forms, especially when anchored in an understanding of the schwa vowel, plays a critical role in developing ESL listening skills. Failure to perceive reduced grammatical elements can distort sentence meaning despite accurate recognition of content words. The improvement observed in this study suggests that raising awareness of schwa as a unifying phonological feature enables learners to adopt more flexible and realistic listening strategies (see Appendix H).

4.4.7 Contraction Processing: From Orthographic Dependence to Perceptual Mapping

Tasks focusing on contractions revealed that learners initially relied heavily on written-form expectations when processing spoken input. In the pre-intervention phase, contracted sequences were often expanded into full forms or segmented into multiple lexical items. Complex contractions such as *shouldn't've gone* were frequently restructured in ways that reflected spelling-based assumptions rather than spoken patterns, indicating an orthography-driven listening approach.

After instruction, learners showed a clear shift in how they processed contracted forms. Post-intervention responses demonstrated greater acceptance of compressed sequences such as *it's*, *I'm*, *you're*, and *didn't*, and an improved ability to map reduced auditory input onto complete syntactic constructions (see Appendix H). Although spelling accuracy remained inconsistent and some learners continued to expand contractions in writing, the reconstructed utterances were largely syntactically and semantically appropriate. This suggests progress in perceptual mapping rather than orthographic control.

These findings reinforce the view that gains in listening comprehension should not be equated with spelling accuracy. In this study, learners' ability to recover intended grammatical meaning from reduced speech proved to be a more meaningful indicator of perceptual development. Even when written forms were imperfect, learners demonstrated increasing sensitivity to how spoken English compresses function words and auxiliaries, reflecting a gradual shift from text-based listening toward speech-oriented processing.

4.4.8 Learner Engagement, Awareness, and Strategy Shift

One of the most notable outcomes of the intervention was a qualitative shift in learners' listening behavior. Learners became more attentive to auditory patterns, more tolerant of ambiguity, and more willing to hypothesize meaning rather than abandoning comprehension attempts. This was evident in their increased tendency to produce phonologically plausible forms instead of unrelated lexical guesses.

This shift aligns with the study's emphasis on perceptual awareness rather than rote performance. Learners were not merely arriving at correct answers but were developing more effective listening strategies. In this sense, the intervention supported not only improved comprehension outcomes but also a fundamental change in how learners approached spoken English.

4.4.9 Educational Implications

The findings suggest that instruction in connected speech should not be treated as an advanced or optional component of ESL curricula. Instead, it should be integrated into listening instruction from the early stages of language learning. In the Pakistani ESL context, classroom instruction has traditionally prioritized grammar, vocabulary, and written accuracy, often leaving learners underprepared for natural, fluent spoken English.

The present study demonstrates that even brief and focused instructional interventions can lead to meaningful changes in learners' perceptual strategies. On this basis, ESL instructors are encouraged to incorporate connected speech features explicitly and systematically into listening lessons. This includes presenting phonological processes in accessible ways, using varied audio materials at manageable speech rates, and designing tasks that require learners to reconstruct meaning from reduced or compressed speech.

Greater emphasis should be placed on the underlying phonological logic of spoken language rather than surface-level accuracy alone. Importantly, partial or imperfect understanding should be recognized as a natural and productive stage in listening development, reflecting learners' gradual adaptation to the realities of authentic spoken communication

Chapter 5

Discussion, Conclusion and Implications

The main purpose of this study was to develop an integrated model of perception of politics with its dual outcomes. In order to attain the purpose, the research questions were articulated and tested with respective hypothesis. Generally, we found a good support for our hypotheses, while some results were contrary to expectation. In following chapter, the possible reasons for these results are discussed in detail.

5.1 Introduction

This chapter interprets the findings of the present study, which examined the impact of explicit, spiral-based instruction in connected speech on the listening comprehension of tertiary-level ESL learners. Rather than simply reiterating the statistical outcomes, the discussion contextualizes the findings by aligning them with the research questions, the guiding theoretical framework, and relevant literature in second language acquisition and applied phonology.

The primary objective of the study was to determine whether systematic instruction in selected connected speech features could enhance learners' ability to perceive and decode naturally occurring spoken English, particularly as encountered through English-language media. Because connected speech phenomena obscure word

boundaries and reduce acoustic salience, they present persistent challenges for ESL learners whose listening instruction often relies on citation-form speech. This chapter therefore explains how and why the instructional intervention contributed to measurable changes in listening performance.

The discussion is organized according to the research questions, with particular attention to differences between the experimental and control groups and to variation in learners' performance across individual connected speech features. The findings are interpreted through established theoretical perspectives, including bottom-up processing theory, cognitive load theory, Krashen's input hypothesis, and Schmidt's noticing hypothesis. The chapter also outlines pedagogical and curricular implications, acknowledges limitations, and offers recommendations for future research.

5.2 Summary of Key Findings

The findings clearly demonstrate that explicit, spiral-based instruction in connected speech significantly enhanced ESL learners' listening comprehension. Comparison of pre-test and post-test results revealed substantial gains in the experimental group across all instructional features, whereas the control group showed only marginal improvement. This contrast suggests that perceptual development in listening does not reliably occur through exposure alone and benefits from systematic instructional support.

Feature-specific analysis indicated differential degrees of improvement. The greatest gains were observed in elision and assimilation, suggesting that once learners became aware of predictable phonological patterns, they were better able to interpret reduced forms in fluent speech. Contractions also showed marked improvement, likely due to learners' prior familiarity with these forms. Weak forms improved more gradually, indicating that perceptual mastery requires sustained exposure and reinforcement. Alveolar flapping remained the most challenging feature, highlighting the influence of L1 phonological systems on perceptual acquisition.

Overall, the findings suggest that ESL listening difficulties stem less from deficiencies in vocabulary or grammar and more from limited perceptual awareness of connected speech phenomena. Targeted instruction and structured practice enabled learners to process authentic spoken language more effectively, laying the foundation for improved listening comprehension.

5.3 Discussion of Findings by Research Question

5.3.1 Effect of Instruction on Overall Listening Comprehension

The first research question examined whether explicit, spiral-based instruction in connected speech influenced learners' overall listening comprehension. The results indicate a significant positive effect, as evidenced by substantial post-test gains in the experimental group compared to the control group.

These improvements suggest that learners' listening difficulties are closely linked to challenges in identifying reduced and modified forms in fluent speech rather than to insufficient lexical or grammatical knowledge. Prior to instruction, learners tended to rely on citation-form expectations, which hindered effective segmentation of continuous speech. Following instruction, learners demonstrated improved decoding ability, allowing them to recognize words and phrases despite phonological reduction and variation.

In contrast, the control group showed no meaningful improvement, indicating that exposure to spoken English alone is insufficient for perceptual development when listening instruction does not explicitly address differences between spoken and written forms.

From a cognitive perspective, the effectiveness of the intervention can be attributed to its spiral-based design, which introduced connected speech features incrementally and reinforced them across instructional cycles. This sequencing reduced cognitive load by enabling learners to focus on recurring phonological patterns, thereby facilitating more efficient processing of spoken input.

These findings support the view that explicit instruction in connected speech strengthens bottom-up processing skills essential for effective listening comprehension and helps bridge the gap between classroom listening activities and authentic spoken language.

The findings indicate that learners' perceptual development trajectories changed following instruction, with the experimental group showing improvement across all targeted connected speech features. However, the extent and rate of improvement varied, reflecting differences in inherent phonological complexity, feature salience, and learners' prior exposure and readiness.

i. Assimilation demonstrated moderate to substantial gains in post-test performance. Prior to instruction, learners struggled to recognize assimilated forms because phonological changes altered segmental realization, obscuring word boundaries. Explicit instruction sensitized learners to predictable phonological variation, enabling more accurate segmentation of continuous speech. These findings suggest that assimilation becomes perceptually accessible when learners recognize it as a systematic rather than arbitrary process.

ii. Elision produced the most pronounced improvement. Once learners understood that certain sounds are routinely omitted in fluent speech, they were better able to reconstruct intended lexical forms. Despite appearing structurally complex, elision proved cognitively accessible when learners were trained to anticipate omission patterns and rely on contextual cues. This supports the view that perceptual difficulty often arises from learners' expectation of complete phonological forms rather than from the absence of segments themselves.

iii. Contractions showed rapid improvement, likely due to learners' existing familiarity with these forms in both spoken and written English. Instruction capitalized on this prior knowledge by aligning orthographic and phonological representations, which facilitated quicker uptake. For this reason, contractions were introduced early in the instructional sequence and functioned as an accessible entry point into connected speech awareness.

iv. Weak forms improved more gradually than other features. Although learners became more sensitive to reduced vowels and stress patterns, fast or dense speech continued to pose perceptual challenges. These findings indicate that mastery of weak forms requires extended exposure and repeated reinforcement, as their low salience makes them less perceptually robust in fluent discourse.

v. Alveolar flapping remained the most challenging feature, even after explicit instruction. Learners frequently confused flap realizations with /r/, likely due to L1 phonological influence and the absence of comparable sounds in their native language. Targeted practice contrasting both sounds e.g., butter yielded limited improvement, underscoring the persistent effect of first-language interference and suggesting that certain phonological features may require prolonged and focused perceptual training.

Collectively, these findings demonstrate that listening difficulties are not uniform across connected speech features but are shaped by phonological salience, frequency of exposure, and cross-linguistic similarity. The results highlight the need for differentiated instructional approaches that account for varying levels of perceptual complexity and allocate instructional time accordingly.

5.4 Relationship of Findings to the Theoretical Framework

The findings provide empirical support for the theoretical frameworks underpinning the study, including bottom-up processing theory, cognitive load theory, Krashen's input hypothesis, and Schmidt's noticing hypothesis. Together, these perspectives explain the observed gains in listening comprehension following explicit connected speech instruction.

From a bottom-up processing perspective, improved post-test performance reflects enhanced ability to decode phonological input. Explicit instruction enabled learners to attend to reduced and altered forms, facilitating more accurate segmentation and integration of phonological and higher-level interpretive processes.

Cognitive load theory helps explain the effectiveness of the spiral-based instructional approach. By introducing one feature at a time and systematically recycling previously taught material, the intervention minimized cognitive overload and promoted gradual automation of perceptual skills.

The findings also align with Krashen's input hypothesis, as instruction increased the comprehensibility of authentic spoken input by making learners aware of phonological patterns that had previously impeded understanding. Similarly, the results support Schmidt's noticing hypothesis, demonstrating that explicit attention to connected speech features transformed previously unnoticed phonological phenomena into meaningful perceptual cues. Collectively, these frameworks function synergistically to explain how explicit instruction enhances listening comprehension by improving perceptual decoding, managing cognitive demands, and directing learner attention.

The findings of this study have important implications for ESL listening instruction, particularly at the postsecondary level where learners are increasingly exposed to authentic spoken English. The demonstrated effectiveness of explicit, spiral-based instruction in connected speech suggests that traditional listening pedagogy, often focused primarily on comprehension questions, does not sufficiently address the perceptual demands of real-world listening.

First, ESL listening instruction should incorporate explicit training in connected speech features rather than relying on exposure alone. The substantial gains observed in the experimental group indicate that difficulties arising from reduced and modified speech forms persist without systematic intervention. Instruction should therefore foreground features such as weak forms, contractions, assimilation, and elision instead of assuming that learners will acquire them implicitly.

Second, the success of the spiral-based instructional model highlights the pedagogical value of sequencing and recycling. Gradual introduction of connected speech features, coupled with repeated reinforcement across instructional cycles, enabled learners to consolidate perceptual awareness without cognitive overload. This approach is particularly beneficial for learners whose prior engagement with

English has been predominantly text-based and who require structured support in adapting to spoken language.

Third, the findings suggest that listening pedagogy should prioritize perceptual development alongside comprehension. Learners' improvement was attributable not only to task practice but also to directed attention to phonetic realization in fluent speech. Activities such as focused listening, guided transcription, minimal-pair discrimination, and awareness-raising tasks can help establish a perceptual foundation that supports more effective comprehension.

The study also underscores the importance of teacher knowledge in implementing connected speech instruction effectively. ESL teachers require a sound understanding of spoken phonology and its pedagogical applications. Accordingly, professional development initiatives should include focused training on teaching connected speech and designing perception-oriented listening activities. These implications are particularly relevant in the Pakistani tertiary context, where listening instruction is often assessment-driven and text-oriented. Integrating explicit connected speech instruction can help narrow the gap between classroom listening practices and the communicative demands of academic and real-life contexts.

5.5 Implications for English as a Second Language Curriculum and Assessment

Beyond classroom pedagogy, the findings have broader implications for ESL curriculum design and listening assessment. The positive impact of explicit connected speech instruction suggests that curricula emphasizing grammatical accuracy and lexical knowledge alone may inadequately prepare learners for authentic spoken communication. At the curricular level, listening syllabi should move beyond exposure-based models and explicitly incorporate connected speech features as learning objectives. Listening should be treated as an active perceptual skill that requires systematic training in decoding fluent speech. A spiral curriculum that introduces connected speech features at progressively higher levels of complexity can support sustained development and reduce long-term perceptual difficulties.

In terms of materials development, ESL textbooks and audio resources should reflect the characteristics of natural spoken English, including reduced forms, variable stress patterns, and phonological linking. Overreliance on carefully articulated speech may inadvertently raise listening expectations and limit learners' preparedness for authentic input. The findings support the inclusion of graded authentic audio materials that increase in phonological complexity while remaining pedagogically accessible.

Listening assessment practices also warrant reconsideration. Conventional assessments tend to emphasize comprehension outcomes while overlooking the perceptual processes underlying successful listening. Incorporating tasks that require recognition of reduced forms, identification of connected speech features, or focused decoding can yield a more accurate representation of learners' listening ability and enhance alignment between instructional objectives and assessment practices.

The study further suggests that connected speech instruction can be integrated progressively across educational levels. At the primary level, instruction may focus on high-frequency and perceptually salient features such as weak forms, basic contractions, and simple linking. At the secondary level, learners may be introduced to more complex patterns such as assimilation and common elision. At the tertiary level, explicit awareness-raising, transcription tasks, and perception-focused activities can refine learners' processing of fluent speech. This progression is intended as an illustrative framework rather than a prescriptive curriculum, with implementation decisions remaining the responsibility of curriculum designers and policymakers.

Effective curricular integration, however, depends on teacher preparation and institutional support. In Pakistan, English teachers are not uniformly trained in phonology or listening pedagogy, and specialized TEFL or TESOL related qualifications are not mandatory. Consequently, large-scale curricular adoption of connected speech instruction would require parallel investment in teacher training and pedagogical guidance.

At the same time, connected speech instruction need not rely exclusively on teacher expertise. Self-access components can supplement classroom instruction by

embedding connected speech activities within textbooks or digital materials. For example, QR codes linking to curated audio resources and guided listening tasks could provide learners with independent exposure to authentic spoken input. Given the widespread availability of mobile devices in Pakistan, mobile-assisted learning offers a practical means of extending perceptual training beyond the classroom. Overall, embedding connected speech awareness into curricular and assessment frameworks can better support learners' transition from classroom-based listening to real-world communicative contexts and promote more realistic, sustainable listening proficiency.

To operationalize these curricular recommendations, a short, structured micro-level instructional sequence can be embedded within regular ESL classes without disrupting existing syllabi. For instance, a fifteen-minute listening segment may begin with a brief awareness-raising phase (three to four minutes), in which learners are introduced to a specific connected speech feature (e.g., weak forms or linking) through isolated examples. This can be followed by a guided perception task (five to six minutes), where students listen to short, authentic or semi-authentic audio clips and identify or discriminate the target feature, thereby strengthening bottom-up processing. Subsequently, a focused decoding activity (three to four minutes), such as partial dictation or micro-transcription, can encourage learners to map acoustic input onto linguistic forms. The segment may conclude with a brief reflective or feedback stage (two to three minutes), allowing learners to verify responses and consolidate noticing. Such a sequence aligns with principles of scaffolded instruction and cognitive load management, as it limits task complexity while systematically building perceptual accuracy. Importantly, its brevity and modularity make it feasible for integration into tertiary-level ESL classrooms as a recurring instructional routine rather than an additional curricular burden.

5.6 Limitations of the Study

Despite the proven efficacy of explicit connected speech instruction, several limitations must be recognized in the interpretation of the results. Acknowledging

these limitations aids in contextualizing the findings and pinpointing areas for enhancement in subsequent research.

First, the study had a small sample size that came from only one tertiary-level institution. The quasi-experimental design facilitated a significant comparison between the control and experimental groups; however, the limited sample constrains the generalizability of the results to larger ESL populations. Conducting the study again with larger and more varied samples would make the results more valid outside of the study.

Second, the length of the teaching intervention was short. While substantial improvements were noted during the intervention period, prolonged instructional exposure is required to attain comprehensive mastery of specific connected speech characteristics, especially those that consistently posed challenges, such as weak forms and alveolar flapping. A prolonged treatment duration may elucidate the sustainability and enduring retention of perceptual improvements.

Third, the study concentrated on a specific subset of connected speech features, including weak forms, contractions, assimilation, elision, and alveolar flapping. Other features of connected speech, like linking, intonation, and rhythm, were not looked at. Consequently, the results provide merely a partial representation of learners' engagement with connected speech phenomena. Even with these problems, the study gives a systematic and theoretically sound look at teaching connected speech and sets the stage for more research into ESL listening pedagogy.

5.7 Recommendations for Future Research

While the present study provides evidence for the effectiveness of explicit connected speech instruction, it also opens several avenues for further investigation. Future research would benefit from involving larger and more diverse samples drawn from multiple institutions and proficiency levels. Such studies would strengthen the generalizability of the findings and allow for meaningful comparisons across different learning contexts. Including learners from a wider range of linguistic

backgrounds may also help clarify how first-language phonological systems shape the perception of connected speech features.

Longitudinal research is particularly needed to examine the long-term impact of explicit connected speech instruction. Extending the duration of instructional interventions and incorporating delayed post-tests could offer insight into whether perceptual gains are maintained over time and how effectively they transfer to spontaneous listening in real-world communicative situations.

Further studies could also broaden the scope of investigation by examining additional aspects of connected speech, such as linking, stress timing, rhythm, and intonation. Exploring these features would contribute to a more comprehensive understanding of spoken language perception. In addition, comparative research contrasting explicit perceptual instruction with implicit exposure-based approaches could help identify the most effective instructional practices for developing listening skills.

Finally, future research may explore the relationship between listening perception and spoken production, particularly how increased perceptual awareness of connected speech influences learners' pronunciation and fluency. Investigating this perception–production relationship could offer valuable insights for designing more integrated and holistic ESL instruction.

To facilitate replication, future studies should adopt a clearly structured implementation of the instructional model by maintaining consistency in key design elements, including participant proficiency level, instructional duration, and task sequencing. The present study can be replicated by employing a controlled pre-test/post-test design with comparable ESL learners, using parallel listening assessment instruments that measure both perceptual discrimination and comprehension. Researchers are encouraged to retain the core instructional sequence comprising awareness-raising, guided perception, and focused decoding tasks while adapting content to their specific linguistic contexts. Instructional materials, including audio samples and task formats, should reflect authentic or semi-authentic connected speech to preserve ecological validity. Additionally, maintaining similar intervention intensity (e.g., short, regular sessions distributed over several weeks) would allow for meaningful

comparison of results. While contextual adaptations are expected, preserving these core methodological features will enable researchers to evaluate the effectiveness of the model across diverse settings and contribute to a more cumulative and comparable body of research.

Bibliography

Baumgardner, R. J. (1993). *The English language in Pakistan*. Oxford University Press.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

Brown, J. D. and Hilferty, A. (2006). Listening for reduced forms. *TESOL Quarterly*, 40(4):835–841.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31(1):21–32.

Cauldwell, R. (2013). *Phonology for listening: Teaching the stream of speech*. Speech in Action.

Celce-Murcia, M., Brinton, D. M., and Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide*. Cambridge University Press, 2 edition.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. MIT Press.

Field, J. (2008). *Listening in the language classroom*. Cambridge University Press.

Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1):55–75.

Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34(2):165–182.

Hamada, Y. (2011). Improvement of listening comprehension skills through shadowing with difficult materials. *Journal of Asia TEFL*, 8(1):139–162.

Hassan, A. (2017). Teaching pronunciation in pakistani esl classrooms: Issues and challenges. *International Journal of English Linguistics*, 7(3):1–9.

Hussain, M. and Javed, N. (2022). Use of listening strategies by pakistani esl learners. *Migration Letters*, 19(Special Issue):120–131.

Iqbal, M. and Ahmed, S. (2021). Effects of pronunciation instruction on esl learners' listening comprehension in pakistan. *Journal of Research in Applied Linguistics*, 12(1):23–38.

Khan, R. and Mahmood, A. (2020). Listening comprehension practices in pakistani esl classrooms: A critical analysis. *Journal of Education and Educational Development*, 7(2):112–128.

Lambert, S. (1992). Shadowing. *Meta*, 37(2):263–273.

Mahmood, M. A. (2018). Teaching listening skills in pakistani esl classrooms: Practices and challenges. *Pakistan Journal of Education*, 35(2):1–15.

Murphey, T. (2001). Exploring conversational shadowing. *Language Teaching Research*, 5(2):128–155.

Nawaz, S. (2021). Listening comprehension problems, corresponding factors, and strategies for better listening skills among esl learners in pakistan. *Pakistan Languages and Humanities Review*, 5(2):1–12.

Pathan, H. (2023). Challenges in esl listening comprehension: Learners' perceptions in higher education. *Journal of Applied Linguistics and Language Research*, 10(2):91–105.

Rahman, F. and Khan, I. (2020). Learners' perceptions of esl listening difficulties at tertiary level in pakistan. *International Journal of English Research*, 6(4):12–19.

Rahman, T. (1990). Pakistani english: The linguistic description of a non-native variety of english. *World Englishes*, 9(1):1–17.

Rattansak, S. (2025). Phonological processes in english connected speech: Implications for l2 speech learning and communication. *Cogent Education*, 12(1):2472474.

Riaz, N., Fatima, N., and Aamer, A. (2023). Enhancing listening skills in pakistani esl classrooms: Challenges and pedagogical implications. *Kurdish Studies*, 11(2):145–160.

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

Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. Continuum.

Rost, M. (2011). *Teaching and researching listening*. Pearson Education, 2 edition.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In Spiro, R. J., Bruce, B. C., and Brewer, W. F., editors, *Theoretical issues in reading comprehension*, pages 33–58. Lawrence Erlbaum Associates.

Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? *Language Teaching*, 56(4):522–538.

Shahid, S., Farukh, A., and Nazar, N. (2021). A study of listening comprehension problems faced by esl learners at undergraduate level in south punjab. *Pakistan Journal of Society, Education and Language*, 7(1):243–253.

Shahzad, S. and Ali, Z. (2019). The role of phonological awareness in improving esl learners' listening comprehension. *Pakistan Journal of Social Sciences*, 39(2):451–462.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2):257–285.

Vandergrift, L. and Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Waqas, M., Anwar, B., and Latif, S. (2025). Teaching weak forms and connected speech in esl pronunciation classrooms. *Asian Journal of English Language Teaching*, 15(1):44–59.

Appendix A

Sample Lesson Plan

A.1 TP1: Lesson Plan of Contractions

Course: ESL Listening and Pronunciation **Level:** Intermediate (B1–B2) **Age Group:** Adult learners

Lesson Length: 60 minutes **Teaching Context:** Tertiary-level ESL classroom **Lesson Aim** To raise learners' awareness of connected speech and its impact on listening comprehension in natural spoken English.

Subsidiary Aims

- To help learners identify reduced forms in continuous speech
- To improve learners' ability to recognize word boundaries during listening

Assumptions Learners are familiar with individual word pronunciation but struggle to recognize words in natural, fast speech. Anticipated Problems and Solutions

Problem Solution Learners may not notice reductions Replay audio with guided focus questions Learners may feel discouraged Use short, manageable listening segments

Materials Audio recording of a short dialogue, Transcript with marked reductions, Whiteboard

Stage and aim	Interaction pattern	Time	Procedure	Trainer's comments
Lead-in: To capture students' interest and introduce the lesson topic, activating prior knowledge and preparing them for the main activities.	T-→ ←S	5 MIN	Introduce the concept of contractions; activate prior knowledge	Engagement; activate schemata
Pre-teaching Vocabulary: To familiarize learners with the lexical items and auxiliary/modal verbs that form the basis of the target contractions, ensuring they can identify and process them in connected speech.	T→ ←S	10 MIN	• Teach words learners will need to understand the audio and do the tasks successfully. • Play an audio containing contractions and ask students to identify what they hear. Response shall be elicited from the students. • Monitor their responses and provide feedback as needed to ensure correct perception and understanding of connected speech features.	1. Provide clear written forms alongside their full expansions (e.g., *I'm = I am*) 2. Use simple example sentences for context 3. Highlight pronunciation differences between formal written forms and spoken contractions 4. Keep the vocabulary list **manageable**, introducing 5–7 items per session to avoid cognitive overload
Listening Practice:	Group work	15 MIN	Play audio containing target contractions. Play the recording; ask learners to identify contractions	Perception focus; connected speech awareness of " I'm, they'll, I'll, They've, It's, She'll, don't, You're"
Noticing / Concept Checking:	T→ ←S	10 MIN	Inductive discovery. Check understanding of meaning and function. Guide learners to discover rules for forming contractions. Ask concept-checking questions (CQS) for each contraction. Students respond to CQS	Ensure comprehension of meaning and grammatical function
Guided Practice / Drilling	T→ ←S	10 MIN	Drill pronunciation of contractions. Shadowing Practice: Model pronunciation and correct errors. Students repeat chorally and individually	Focus on connected speech: reduction
Contextualized Practice	Group work	10 MIN	Practice contractions in sentences. Provide sentences, dialogues, or song excerpts; guide practice. Students read aloud and repeat.	Apply contractions in context; reinforce listening and speaking
Monitoring and Feedback	Throughout		Observe learners during activities. Monitor responses; provide corrective feedback on pronunciation and comprehension. Respond to feedback; adjust pronunciation.	Continuous support; reinforce correct perception

FIGURE A.1: Instructional Procedure for Teaching Contractions Stage 1

Stage and aim	Interaction pattern	Time	Procedure	Trainer's comments
Lead-in: To activate prior knowledge, reinforce the concepts, and prepare them for the main activities.	T-→ ←S	5 MIN	Activate prior knowledge	Engagement; activate schemata
Pre-teaching Vocabulary: To familiarize learners with the lexical items and auxiliary/modal verbs that form the basis of the target contractions, ensuring they can identify and process them in connected speech.	T→ ←S	10 MIN	• Teach words that the learners will need to understand the audio and do the tasks successfully. • Play an audio containing contractions and ask students to identify what they hear. • Monitor their responses and provide feedback as needed to ensure correct perception and understanding of connected speech features.	• Provide clear written forms alongside their full expansions (e.g., you'd = *You would/had*) • Use simple example sentences for context • Highlight pronunciation differences between formal written forms and spoken contractions • Keep the vocabulary list **manageable**, introducing 8 items per session to avoid cognitive overload
Listening Practice:	**Group work**	15 MIN	Play audio containing target contractions. Play the recording; ask learners to identify contractions. Response shall be elicited from the students	Perception focus; connected speech awareness of "shouldn't, wouldn't, couldn't, I'd, You'd, They'd, didn't, we'd"
Noticing / Concept Checking:	T→ ←S	10 MIN	Check understanding of meaning and function. Ask concept-checking questions (CQS) for each contraction. Students respond to CQS	Ensure comprehension of meaning and grammatical function
Guided Practice / Drilling	T→ ←S	10 MIN	Drill pronunciation of contractions. Model pronunciation and correct errors. Students repeat chorally and individually	Focus on connected speech: reduction
Contextualized Practice	Group work	10 MIN	Practice contractions in sentences. Provide sentences, dialogues, or song excerpts; guide practice. Students read aloud and repeat.	Apply contractions in context; reinforce listening and speaking
Monitoring and Feedback	Throughout		Observe learners during activities. Monitor responses; provide corrective feedback on pronunciation and comprehension. Respond to feedback; adjust pronunciation.	Continuous support; reinforce correct perception

FIGURE A.2: Stage 2: This stage includes contractions that are easy to miss in connected speech.

Stage and aim	Interaction pattern	Time	Procedure	Trainer's comments
Lead-in: Activating prior knowledge and preparing them for the main activities.	T-→ ←S	5 MIN	Activate prior knowledge	Engagement; activate schemata
Pre-teaching Vocabulary: To familiarize learners with the lexical items and auxiliary/modal verbs that form the basis of the target contractions, ensuring they can identify and process them in connected speech.	T→ ←S	10 MIN	• Teach words learners will need to understand the audio and do the tasks successfully. • Play an audio containing contractions and ask students to identify what they hear. • Monitor their responses and provide feedback as needed to ensure correct perception and understanding of connected speech features.	• Provide clear written forms alongside their full expansions (e.g., *I'd've = I would have*) • Use simple example sentences for context • Highlight pronunciation differences between formal written forms and spoken contractions • Keep the vocabulary list **manageable**, introducing 8 items per session to avoid cognitive overload.
Listening Practice:	Group work	15 MIN	Play audio containing target contractions. Play the recording; ask learners to identify contractions. Response shall be elicited from the students.	Perception focus; connected speech awareness of "I'd've, ain't, shouldn't've, gonna, wanna, gotta, lemme, sorta"
Noticing / Concept Checking:	T→ ←S	10 MIN	Check understanding of meaning and function. Ask concept-checking questions (CQS) for each contraction. Students respond to CQS	Ensure comprehension of meaning and grammatical function
Guided Practice / Drilling	T→ ←S	10 MIN	Drill pronunciation of contractions. Model pronunciation and correct errors. Students repeat chorally and individually	Focus on connected speech: reduction
Contextualized Practice	Group work	10 MIN	Practice contractions in sentences. Provide sentences, dialogues, or song excerpts; guide practice. Students read aloud and repeat.	Apply contractions in context; reinforce listening and speaking
Monitoring and Feedback	Throughout		Observe learners during activities. Monitor responses; provide corrective feedback on pronunciation and comprehension. Respond to feedback; adjust pronunciation.	Continuous support; reinforce correct perception

FIGURE A.3: Stage 3: In this session more advanced contractions are introduced along with their full forms.

Appendix B

Miscellaneous Listening Activities and Drills

Listening Exercise 1 Topic: Contractions

Instruction: Listen to the audio and fill in the blanks.

1. .....going to visit my friend today.
2. ..... it crazy how this works?
3. .....be here in five minutes.
4. Pardon me. ..... sit elsewhere.
5. .....not sure if .....a good idea.
6. .....been waiting for an hour.
7. They ..... told us about the gifts.
8. .....been curious if you .....tell me.
9. ..... getting a little cold.
10. He said ..... be right back.
11. You ..... told me.

12. ..... better finish your work.

13. ..... in pain.

Listening Exercise 2 Topic: Contractions

Instruction: Listen to the audio and fill in the gaps.

Interviewer 1: But seriously, just talking about the season, it is interesting, because I know all of you thought about what do you do when you do have to give someone bad news. I know some of you are less hesitant than others, but is it hard to give someone bad news early on?

Judge 1: Well, not hard, hard in the fact that yes, going to be the truth. Now, on live television, and now we have to tell you and critique your entire future.

I think of it as my kids. going to be as dignified as I possibly be, but at the same time, ... going to tell you if ..... not ready, ... not ready.

Interviewer 1: Yeah.

Judge 1: Because, remember now, ... our reputation and their future.

Interviewer 1: When you look at what chosen to bring back to Hollywood, are you happy with it?

Judge 2: Oh my gosh. Oh yeah.

Interviewer 1: I saw what happened yesterday. You must be psyched.

Judge 2: I think ..... so thrilled. ..... put ourselves into a pickle, honestly.

Interviewer 2: Is it because . all been there? I mean, in seriousness, . all had a moment where maybe it ..... go your way.

Listening Exercise 3

Instruction: Identify the use of contractions in the dialogue.

Rosalie: I ..... particularly like you, but... Well, I envy you.

Bella: What? ..... ridiculous.

Rosalie: No, not. You have a choice. I . None of us did, but you do, and choosing wrong. I ..... care how miserable your human life is.

Bella: My life is not miserable. And not perfect. I mean, nobody's life is perfect.

Rosalie: Mine was. Absolutely perfect. There were things I still wanted. A family of my own.

Bella: ..... nothing ..... ever going to want. More.

Rosalie: wrong again. After been changed, one thing want more. One thing kill for. Blood.

Appendix C

Research Instruments

C.1 Audio Source

The audio materials for both the pre-test and post-test were scripted in advance and recorded by a native speaker of American English to ensure consistency, clarity, and reliability across test administrations. The speaker provided informed consent for the use of the recordings for research purpose.

Pre-Test

Connected Speech Aspects		Phonetic Transcription
Type	**Phrase**	**IPA (GA)**
1. Progressive Assimilation	dogs bark	/dɔgz bark/ → [dɔgzbark]
2. Regressive (Assimilation	input	/ɪnpʊt/ → [ɪmpʊt]
3. Coalescent Assimilation	did you	/dɪd ju/ → [dɪdʒə] or [ˈdɪdʒu]
4. Place Assimilation	ten boys	/tɛn bɔɪz/ → [tɛmbɔɪz]
5. Manner Assimilation	that side	/ðæt saɪd/ → [ðætsaɪd]
6. Voicing Assimilation	have to go	/hæv tu goʊ/ → [hæf tə goʊ]
7. Partial Assimilation	Has to	/hæz tu/ → [hæs tə]
8. Complete Assimilation	good girl	/gʊd gɝl/ → [gʊgɝl]
9. Consonant Elision	next day	/nɛkst deɪ/ → [nɛksdeɪ]
10. Vowel Elision	library	/ˈlaɪbri/
11. Initial Elision	'cause (because)	/bɪˈkɔz/ → [kəz] or [kʌz]
12. Final Elision	first time	/fɝst taɪm/ → [fɝstaɪm]
13. Across-Word Elision	last night	/læst naɪt/ → [læsnaɪt]
14. Intervocalic Flapping	water	/ˈwɔtər/ → [ˈwarə]
15. Flapping After /r/	party	/ˈpartɪ/ → [ˈparɹi]
16. Flapping Across Words	get out	/gɛt aʊt/ → [gɛr aʊt]
17. Nasal Flapping	butter	/ˈbʌtər/ → [ˈbʌɾə]
18. Voiceless vs. Voiced Flaps	ladder	/ˈlædər/ → [ˈlærə]
19. Weak Form (Articles)	a cat	/ei kæt/ → [əkæt]
20. Weak Form (Prepositions)	to school	/tu skuːl/ → [tə skuːl]
21. Weak Form (Pronouns)	tell him	/tɛl hɪm/ → [tɛlɪm]
22. Weak Form (Aux/Modal)	I can go	/aɪ kæn goʊ/ → [aɪ kən goʊ]
23. Weak Form (Conjunctions)	bread and butter	/brɛd ən ˈbʌɾə/ → [brɛd ən ˈbʌɾə]
24. Weak Form (Determiners)	your bag	/jɔr bæg/ → [jə bæg]
25. Weak Form (Miscellaneous)	been there	/bɪːn ðɛr/ → [bɪnðɛr]
26. Aux + Not Contraction	can't go	/kænt goʊ/ → [kæŋgoʊ]
27. Pronoun + Aux	I'm happy	/aɪ əm ˈhæpi/ → [əm hæpi]
28. Noun + Aux	John's here	/dʒanz hɪər/ → [dʒanz ɪr]
29. Tag/Short Form	isn't it	/ˈɪzənt ɪt/ → [ˈɪzənɪt]
30. Informal Contraction	gonna leave	/ˈgoʊɪŋ tu liv/ → [ˈgʌnə liv]

FIGURE C.1: Pre-Test Script

Post-Test

Subtypes of Connected Speech Aspects		Phonetic Transcription
Type	Phrase	IPA (GA)
1. Progressive Assimilation	dogs biting	/dɔgz baɪtɪŋ/ → [dɔgzˈbaɪrɪŋ]
2. Regressive Assimilation	In case	/ɪn keɪs/ → [ɪŋ keɪs]
3. Coalescent Assimilation	would you	/wʊd ju/ → [ˈwʊdʒu] or [ˈwʊdʒə]
4. Place Assimilation	ten points	/tɛn pɔɪnts/ → [tɛm pɔɪnts]
5. Manner Assimilation	Good night	/gʊd naɪt/ → [gʊ naɪt]
6. Voicing Assimilation	Dogs chase	/dɔgz tʃeɪz/→ [dɔgztʃeɪs]
7. Partial Assimilation	green beans	/grin binz/ → [grim binz]
8. Complete Assimilation	this year	/ðɪs jɪr/ → [ðɪʃ jɪr]
9. Consonant Elision	next week	/nɛkst wik/ → [nɛkswik]
10. Vowel Elision	chocolate bar	/ˈtʃaklət bar/ → [ˈtʃaklɪt bar]
11. Initial Elision	Hit them all	/hɪt ðɛm ɔl/ → [hɪr əm ɔl]
12. Final Elision	Suppose to	/səˈpoʊzd tu/ → [səˈpoʊstə]
13. Across-Word Elision	The orange	/ði ɔrinʤ/ → [ðɔːrɪnʤ]
14. Intervocalic Flapping	water bottle	/ˈwɔtər ˈbatəl/ → [ˈwɔrə ˈbarl]
15. Flapping After /r/	party animal	/ˈparti ˈænɪməl/ → [ˈparri ˈænɪməl]
16. Flapping Across Words	put it away	/pʊt ɪt əˈweɪ/ → [pʊr ɪr əˈweɪ]
17. Nasal Flapping	mountain air	/ˈmaʊntən ɛr/ → [ˈmaʊrn̩ər]
18. Voiceless vs. Voiced Flaps	glitter	/ˈglɪtər/ → [ˈglɪrə]
19. Weak Form (Articles)	a little noise	/ ə ˈlɪtəl nɔɪz /→[ə ˈlɪrl nɔɪz]
20. Weak Form (Prepositions)	from the start	/from ðə start/→[frəm ðə start]
21. Weak Form (Pronouns)	Their shoes	/ðeəʳ ʃuːz /→ [ðər ʃuːz]
22. Weak Form (Aux/Modal)	I can't help it	/ɑɪ kænt hɛlp ɪt/ → [ɑɪ kænʔɛlpɪr]
23. Weak Form (Conjunctions)	salt and pepper	/sɔlt ænd ˈpɛpər/→ [[sɔl ən ˈpɛpɚ]
24. Weak Form (Determiners)	there are some apples	/ðɛr ər səm ˈæpəlz/ → [ðər ər smæpəlz]
25. Weak Form (Miscellaneous)	I'd've been	/ɑɪd æv bɪːn/ → [ɑɪrəbɪn]
26. Aux + Not Contraction	shouldn't have gone	/ˈʃʊdənt hæv gɔn/ → [ˈʃʊdn̩ əv gɔn]
27. Pronoun + Aux	she's been working	/ ʃiːz bɪːn ˈwɜkɪŋ /→ [ʃiz bɪn ˈwɜkɪŋ]
28. Noun + Aux	Sarah is coming	/ ˈsɛɹə iz ˈkʌmɪŋ /→/ˈsɛɹəz ˈkʌmɪŋ/
29. Tag/Short Form	you're coming, aren't you?	/jʊr ˈkʌmɪŋ ˈarnt ju/ → [jə ˈkʌmɪŋ ˈarnʔ jə]
30. Informal Contraction	gotta be there	/ˈgatə bi ðɛr/ → [ˈgarə bi ðər]

FIGURE C.2: Post-Test Script

Code	Assim Avg	Assim Weighted	Elision Avg	Elision Weighted	Flap Avg	Flap Weighted	Weak Avg	Weak Weighted	Con Avg	Con Weighted	Final Score (/100)	Performance Level
Phrases	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!	#REF!
CG1	2.5	10	5	20	3.6	14.4	3.1428571	12.57142857	2.2	8.8	65.77142857	Moderate
CG2	2.125	8.5	5	20	2.2	8.8	2.5714286	10.28571429	2.8	11.2	58.78571429	Developing
CG3	2.75	11	3.8	15.2	1.2	4.8	2.7142857	10.85714286	2.4	9.6	51.45714286	Developing
CG4	2.5	10	4.4	17.6	3	12	2.8571429	11.42857143	2.6	10.4	61.42857143	Moderate
CG5	3.375	13.5	4.2	16.8	4	16	3.4285714	13.71428571	3.2	12.8	72.81428571	Moderate
CG6	2.75	11	4.2	16.8	3	12	2.7142857	10.85714286	3.6	14.4	65.05714286	Moderate
CG7	2	8	3	12	2	8	2.5714286	10.28571429	3.2	12.8	51.08571429	Developing
CG8	4	16	3.8	15.2	4	16	3.2857143	13.14285714	2.2	8.8	69.14285714	Moderate
CG11	3.25	13	4.4	17.6	4.8	19.2	2.8571429	11.42857143	2.8	11.2	72.42857143	Moderate
CG12	4.625	18.5	3.8	15.2	4.2	16.8	2.7142857	10.85714286	4.6	18.4	79.75714286	Good
CG13	2.25	9	4.4	17.6	2	8	2.5714286	10.28571429	3.6	14.4	59.28571429	Developing
CG15	3	12	4.8	19.2	3.6	14.4	3.4285714	13.71428571	3.6	14.4	73.71428571	Moderate
CG16	2.375	9.5	4	16	1.4	5.6	2.7142857	10.85714286	3.6	14.4	56.35714286	Developing
CG17	3.25	13	3.8	15.2	2	8	3	12	3.2	12.8	61	Moderate
CG18	3.125	12.5	3.4	13.6	3	12	3	12	3.2	12.8	62.9	Moderate
CG21	2.75	11	5	20	2.8	11.2	3.1428571	12.57142857	4.2	16.8	71.57142857	Moderate

FIGURE C.3: Control Group's Pre-Test Averaged and Weighted

St Code	Assim Avg	ssim Weighted	Elision Avg	lision Weighted	Flap Avg	Flap Weighted	Weak Avg	Weak Weighted	Con Avg	Con Weighted	Final Score (/100)	erformance Lev
Phrases		0	#REF!	#REF!	#REF!	#REF!					#REF!	#REF!
CG1	2.375	9.5	3.2	12.8	2.8	11.2	3.5714286	14.28571429	2.8	11.2	62.3	Moderate
CG2	2.5	10	3.2	12.8	3	12	4	16	3.2	12.8	57.77142857	Developing
CG3	2.375	9.5	4	16	3	12	3.1428571	12.57142857	2.6	10.4	55.9	Developing
CG4	2.75	11	3.4	13.6	2.4	9.6	3	12	1.6	6.4	61.85714286	Moderate
CG5	3.25	13	4	16	3	12	3.7142857	14.85714286	3.2	12.8	67.28571429	Moderate
CG6	2.875	11.5	3	12	3	12	3.5714286	14.28571429	3	12	57.67142857	Developing
CG7	2.5	10	3.2	12.8	2.2	8.8	3.1428571	12.57142857	2.4	9.6	53.54285714	Developing
CG8	3	12	3.4	13.6	2.8	11.2	3.2857143	13.14285714	2.2	8.8	62.97142857	Moderate
CG11	3.75	15	4.2	16.8	2.8	11.2	4.1428571	16.57142857	2.4	9.6	75	Good
CG12	3.875	15.5	3.2	12.8	3.4	13.6	4	16	4	16	65.9	Moderate
CG13	3.125	12.5	3.4	13.6	2.2	8.8	3	12	3	12	67.58571429	Moderate
CG15	3.375	13.5	3.6	14.4	3.2	12.8	4.5714286	18.28571429	3.6	14.4	60.92857143	Moderate
CG16	2.25	9	3.2	12.8	2.2	8.8	2.8571429	11.42857143	2.2	8.8	56.77142857	Developing
CG17	3.125	12.5	3.4	13.6	2.2	8.8	3.1428571	12.57142857	3.4	13.6	53.98571429	Developing
CG18	2.875	11.5	2.8	11.2	2.2	8.8	2.5714286	10.28571429	2.2	8.8	59.38571429	Developing
CG21	3	12	3.8	15.2	3	12	3.5714286	14.28571429	3.4	13.6	39.2	Beginner

FIGURE C.4: Control Group's Post-Test Averaged and Weighted

) Code	Assim Avg	Assim Weighted	Elision Avg	Elision Weighted	Flap Avg	Flap Weighted	Weak Avg	Weak Weighted	Con Avg	on Weighte	Score (/1	rformance Le
Phrases	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
EG1	3.25	13	4.8	19.2	3.4	13.6	3.571428571	14.28571429	2.2	8.8	68.88571	Moderate
EG2	3.5	14	5	20	4	16	3.428571429	13.71428571	2.4	9.6	73.31429	Moderate
EG3	3.5	14	4.4	17.6	4	16	3.285714286	13.14285714	3.6	14.4	75.14286	Good
EG4	3.25	13	4.4	17.6	4	16	3.285714286	13.14285714	3.6	14.4	74.14286	Moderate
EG5	3.375	13.5	3.6	14.4	4.2	16.8	3.142857143	12.57142857	4	16	73.27143	Moderate
EG6	2.875	11.5	4	16	4.4	17.6	3.428571429	13.71428571	3	12	70.81429	Moderate
EG7	3.375	13.5	3.8	15.2	2.4	9.6	3.142857143	12.57142857	2.2	8.8	59.67143	Developing
EG8	2.5	10	4.4	17.6	4	16	2.285714286	9.142857143	2	8	60.74286	Moderate
EG9	2.375	9.5	3.6	14.4	3.4	13.6	3.142857143	12.57142857	2	8	58.07143	Developing
EG10	1.375	5.5	3.2	12.8	1.6	6.4	1.571428571	6.285714286	1.4	5.6	36.58571	Beginner
EG11	3.375	13.5	4.6	18.4	4.2	16.8	3.857142857	15.42857143	5	20	84.12857	Good
EG12	3.75	15	2.6	10.4	2.2	8.8	2.571428571	10.28571429	1.8	7.2	51.68571	Developing
EG13	3.5	14	3	12	4	16	2.714285714	10.85714286	1.8	7.2	60.05714	Moderate
EG14	3.125	12.5	4.6	18.4	3.2	12.8	2.571428571	10.28571429	2	8	61.98571	Moderate
EG15	3	12	4.6	18.4	1.8	7.2	2.571428571	10.28571429	2.6	10.4	58.28571	Developing
EG16	2.5	10	3.8	15.2	1.8	7.2	2.285714286	9.142857143	2.6	10.4	51.94286	Developing

FIGURE C.5: Experiment Group's Pre-Test Averaged and Weighted

St Code	Assim Avg	Assim Weighted	Elision Avg	Elision Weighted	Flap Avg	Flap Weighted	Weak Avg	Weak Weighted	Con Avg	Con Weighted	Final Score (/100)	ormance Le
Phrases	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
EG1	3.75	15	4.6	18.4	3.8	15.2	3.7142857	14.85714286	4	16	81.85714286	Good
EG2	4	16	4.8	19.2	4.6	18.4	4.2857143	17.14285714	4.6	18.4	89.94285714	Good
EG3	4.25	17	4.8	19.2	5	20	4.1428571	16.57142857	4.8	19.2	91.17142857	Excellent
EG4	4.25	17	4.6	18.4	5	20	4.2857143	17.14285714	4.6	18.4	89.34285714	Good
EG5	3.75	15	3.8	15.2	4.4	17.6	4.2857143	17.14285714	4.2	16.8	77.74285714	Good
EG6	4.25	17	4.4	17.6	3.8	15.2	4.5714286	18.28571429	3.2	12.8	83.28571429	Good
EG7	4.125	16.5	4.8	19.2	3.6	14.4	4.4285714	17.71428571	3.8	15.2	79.81428571	Good
EG8	3.125	12.5	3.4	13.6	2.8	11.2	3.5714286	14.28571429	3	12	67.58571429	Moderate
EG9	4.125	16.5	4.8	19.2	5	20	4.2857143	17.14285714	4	16	88.04285714	Good
EG10	3	12	4	16	4.6	18.4	4	16	3.8	15.2	81.6	Good
EG11	4.375	17.5	4.4	17.6	4.4	17.6	4.5714286	18.28571429	4.8	19.2	89.38571429	Good
EG12	3.25	13	4.6	18.4	4.6	18.4	5	20	4.6	18.4	84.2	Good
EG13	4.5	18	4.4	17.6	3.4	13.6	4	16	3.6	14.4	76.4	Good
EG14	3	12	3.6	14.4	4	16	3.8571429	15.42857143	2.8	11.2	73.82857143	Moderate
EG15	3	12	4.8	19.2	5	20	3.8571429	15.42857143	4	16	81.82857143	Good
EG16	3.375	13.5	4	16	4.4	17.6	4.1428571	16.57142857	3.8	15.2	63.67142857	Moderate

FIGURE C.6: Experiment Group's Post-Test Averaged and Weighted

EG Code	Pre-test Score	Post-test Score	Gain Score
EG10	36.59	81.60	+45.01
EG12	51.69	84.20	+32.51
EG9	58.07	88.04	+29.97
EG15	58.29	81.83	+23.54
EG7	59.67	79.81	+20.14
EG2	73.31	89.94	+16.63
EG13	60.06	76.40	+16.34
EG3	75.14	91.17	+16.03
EG4	74.14	89.34	+15.20
EG1	68.89	81.86	+12.97
EG6	70.81	83.29	+12.47
EG14	61.99	73.83	+11.84
EG8	60.74	67.59	+6.84
EG11	84.13	89.39	+5.26
EG5	73.27	77.74	+4.47

FIGURE C.7: Pre-Test and Post-Test Gain Score of Experimental Group

Appendix D

Audio Scripts and Transcripts

D.1 Song Transcription: "Fix you" by Cold Play

When you try your best, but you don't succeed

When you get what you want, but not what you need

When you feel so tired, but you can't sleep

Stuck in reverse

When the tears come streaming down your face

'Cause you lose something, you can't replace

When you love someone, but it goes to waste

What could it be worse?

Well, high up above, or down below

When you're too in love to let it go

But if you never try, you'll never know

Just what you're worth

D.2 Transcripts of the audios taken from the videos of YouTube that were used in the intervention for listening purpose

Delivery of the Dialogues at Slow Pace: A scene from the movie "The Sixth Sense"

Listen and notice how the highlighted words are spoken.

Mom: Hmm, geez, I hope nobody **got** hurt. You're very quiet. You're mad I missed **the** play, aren't you? I **have** two jobs, baby. You know how important they **are** for us. I'd give anything to have **been** there. **Kid:** I'm ready **to** communicate **with** you now. **Mom:** Communicate? **Kid:** Tell you **my** secrets. **Mom:** What is it? **Kid:** You know the accident up there? **Mom:** Yeah. **Kid:** Someone **got** hurt. **Mom:** They did? **Kid:** A lady. She died. **Mom:** Oh my god, **but you can** see **her**? **Kid:** Yes. **Mom**: **Where is** she? **Kid:** Standing next to my window. **Mom:** Oh, you're scaring me. **Kid:** **They** scare me too sometimes. **Mom:** They?. **Kid:** Ghosts. **Mom:** You see ghosts, Cole? **Kid:** They want me **to** do things for them.	**Mom:** They talk **to** you? They **tell** you **to** do things? **Kid:** They're the ones **that** used **to** hurt me. What **are** you thinking, Mama? You think **I'm** a freak? **Mom:** Look **at** my face. I would never think that **about** you. Ever. Got it? **Kid:** Got it. **Kid: Grandma says** hi. She **says** she's sorry **for** taking the bumblebee pendant. She dislikes it a lot. **Mom**: What? **Kid:** Grandma comes to visit me sometimes. **Mom**: Cole, **that's** very wrong**. Grandma's** gone, you know that. **Kid:** I know. **Kid:** She wanted me to tell you… **Mom:** Cole, please stop. **Kid:** She wanted me **to** tell you she saw you dance. She said when you were little, you **and** her had a fight right before your dance recital. You thought she **didn't** come **to** see you dance? She did. She hid in **the** back so you wouldn't see. She said you were like **an** angel. She said you came **to the** place where they buried her, asked her a question. She said the answer is... "Every day". What did you ask? **Mom:** Do... Do I make her proud?

FIGURE D.1: Dialogues at slow pace (Movie: The Sixth Sense)

Delivery of the Dialogues at Normal Pace: Scene from the movie "Twilight"

Rosalie: I **don't** particularly like you, but... Well, I envy you. **Bella:** What? **That's** ridiculous. **Rosalie:** No, **it's** not. You have a choice. I **didn't**. None of us did, but you do, and **you're** choosing wrong. I **don't** care how miserable your human life is. **Bella:** My life is not miserable. And **it's** not perfect. I mean, **nobody's** life is perfect.	**Rosalie:** Mine was. Absolutely perfect. There were things I still wanted. A family of my own. Royce King was the most eligible bachelor in town **Bella: There's** nothing **I'm** ever going to want. More. **Rosalie: You're** wrong again. After **you've** been changed, **there's** one thing **you'll** want more. One thing **you'll** kill for. Blood.

FIGURE D.2: Dialogues at Normal Pace (Movie: The Twilight

Delivery of the Dialogues at Fast Pace: Backstage talk of American Idol Judges

Interviewer 1: But seriously, just talking about the season, it is interesting, because I know all of you thought about what do you do when you do have to give someone bad news. I know some of you are less hesitant than others, but is that, is it hard to give someone bad news early on? **Judge 1:** Well, **it's** not hard, **it's** hard in the fact that yes, **it's** going to be the truth. Now, **we're** on live, or should I say television at a point, and now we have to tell you and critique your entire future. I think of it as my kids. **I'm** going to be as dignified as I possibly be, but at the same time, **I'm** going to tell you if **you're** not ready, **you're** not ready.	**Inteviewer 1:** Yeah. **Judge 1:** Because, remember now, **it's** our reputation and their future. **Inteviewer 1:** When you look at what **they've** chosen to bring back to Hollywood, are you happy with it? **Judge 2:** Oh my gosh. Oh yeah. **Inteviewer 1:** I saw what happened yesterday You must be psyched. **Judge 2:** I think **we're** so thrilled. **We've** put ourselves into a pickle, honestly. **Interviewer 2:** Is it because **you've** all been there? I mean, in seriousness, **you've** all had a moment where maybe it **didn't** go your way, or you were recording….

FIGURE D.3: Dialogues at Fast Pace (Program: American Idol

Appendix E

Scoring Rubric

Table E.1: Scoring Rubric for Recognition Task

Score	**Descriptor**	**Criteria (Recognition Task)**
5 (Excellent)	Fully correct recognition	The phrase was transcribed exactly as heard, with no errors or omissions. Minor acceptable spelling differences that do not alter meaning were disregarded.
4 (Good)	Minor, non-meaningful error	The response showed clear understanding with one small phonetic or orthographic deviation that did not change the intended words.
3 (Fair)	Partial recognition	Only one word was accurately recognized or the overall phrase was partially correct, indicating developing auditory perception.
2 (Poor)	Related but incorrect	The response bore phonological resemblance to the target phrase but reflected misunderstanding of sound blending or reduction.
1 (Very Poor)	No recognition	The response was blank, irrelevant, or completely incorrect, showing no evidence of connected speech recognition.

Appendix F

Translated Focus Group Discussion

Table F.1: Translated Participant Responses from Post-Module Interviews

Speaker	**Excerpt**
Researcher	The original discussion was in Urdu, so it was translated in English and participants were given codes such as P1, P2 and so on.
Researcher	How was your experience with the listening module?
P1	It was very good. We learned new things, it helped us with pronunciation, and our knowledge of accents improved. This will help us in the future.
Researcher	Do you think your listening skills improved or not?
P1	Yes, our listening improved a lot. When we used to listen to international news, there were contractions and similar features that we couldn't understand. With the help of these lessons, we can now understand those things.
Researcher	How was your experience?

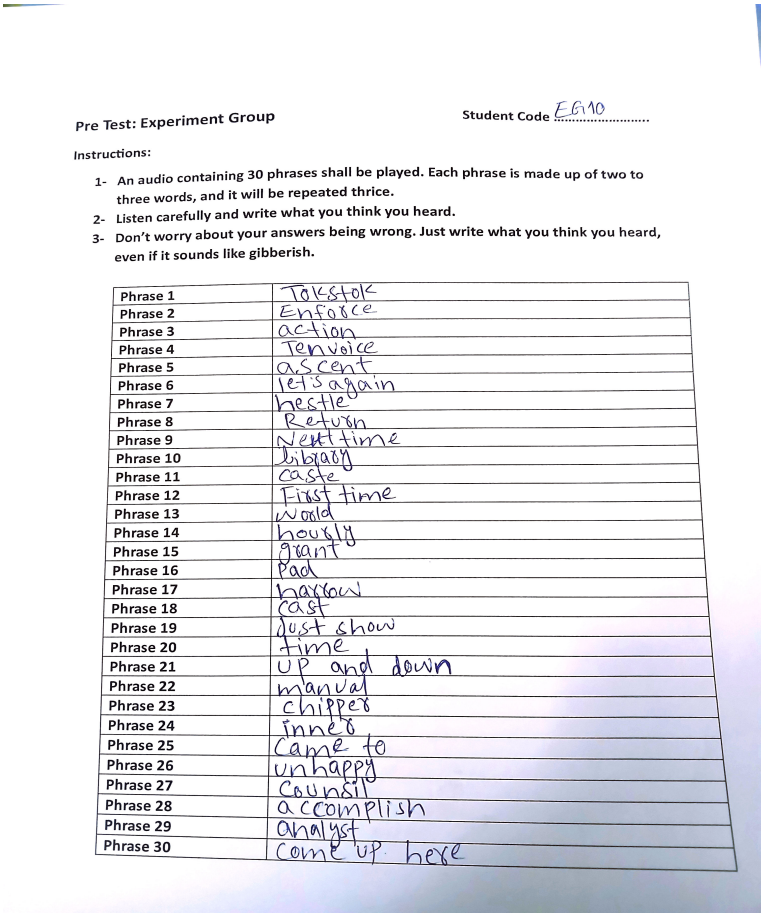
Speaker	**Excerpt**
P2	The experience was very good and enjoyable. The best thing is that our English has improved a lot compared to before. We can speak better, we can listen better, and when someone speaks English, we can understand and respond properly.
Researcher	Please share your experience.
P3	The experience was quite good. We learned many things that we had never been taught before. What we learned here will help us a lot in the future. We used to hear different accents but couldn't understand them clearly; because of this practice, things that were difficult or unfamiliar have now become clear.
Researcher	How was your experience?
P4	It was a very good experience. Our listening improved, and our speaking skills also improved.
P5	The experience was very good. Earlier, when we watched movies, we couldn't understand many words. Now we can understand much more. Our listening has improved a lot, and overall, it was a good experience.
P6	It was a very good experience. Earlier we didn't have confidence or ability to speak, but now we do, and we can speak correctly.
P7	We enjoyed learning with you. The class environment was friendly, which made it easy to understand what you taught. Our knowledge increased, and our listening and speaking improved further.

Speaker	**Excerpt**
P8	Alhamdulillah, our experience was very good. Before these sessions, we couldn't understand foreigners' discussions or English movies. After this, contractions and related features became clear, and now understanding has become easier.
P9	The experience was very good. We learned a lot about communication how to listen, understand others, and respond appropriately.
Researcher	Please share your experience.
P10	Our experience was very good, especially our listening skills improved. Earlier we faced difficulty with pronunciation of words, but with practice, that also improved.
P11	Our experience was quite good. Our understanding developed, and we learned many new things.
P12	It was a very good experience. Our listening skills improved.
P13	I would say that we learned a lot.
Researcher	Do you think this kind of module should be added to regular English classes?
P14	Yes, it should be added. Things that we couldn't understand before start to become clear.
P15	Yes, it is very important, because if we don't understand English properly, we cannot respond correctly.
Researcher	Should this practice be introduced at school level, or is it suitable only for university level?
P16	It should start from the beginning. If we had learned this earlier, our English would be much better by now. It can still improve if we try.
Researcher	Do you think this module helped only listening, or speaking as well?

Speaker	**Excerpt**
P16	It helped both listening and speaking, because when you gain knowledge, you can apply it anywhere.

Appendix G

Data Collection Sample



Pre Test: Experiment Group Student Code EG10

Instructions:

1- An audio containing 30 phrases shall be played. Each phrase is made up of two to three words, and it will be repeated thrice.
2- Listen carefully and write what you think you heard.
3- Don't worry about your answers being wrong. Just write what you think you heard, even if it sounds like gibberish.

Phrase 1	Tokstok
Phrase 2	Enforce
Phrase 3	action
Phrase 4	Tenvoice
Phrase 5	aScent
Phrase 6	let's again
Phrase 7	hestle
Phrase 8	Return
Phrase 9	Next time
Phrase 10	library
Phrase 11	caste
Phrase 12	First time
Phrase 13	world
Phrase 14	hourly
Phrase 15	grant
Phrase 16	Pad
Phrase 17	harbour
Phrase 18	cast
Phrase 19	just show
Phrase 20	time
Phrase 21	up and down
Phrase 22	manual
Phrase 23	chippers
Phrase 24	inner
Phrase 25	Came to
Phrase 26	unhappy
Phrase 27	Council
Phrase 28	accomplish
Phrase 29	analyst
Phrase 30	come up here

FIGURE G.1: Participant EG10 Pre-Test

Post Test: Experiment Group

Student Code EG10

Instructions:

1- An audio containing 30 phrases shall be played. Each phrase is made up of two to three words, and it will be repeated thrice.
2- Listen carefully and write what you think you heard.
3- Don't worry about your answers being wrong. Just write what you think you heard, even if it sounds like gibberish.

Phrase 1	~~Falls pining~~ Dogs fighting
Phrase 2	In case
Phrase 3	reject
Phrase 4	Tempoint
Phrase 5	delight
Phrase 6	Toxis
Phrase 7	~~we need~~ we meet
Phrase 8	This year
Phrase 9	Next week
Phrase 10	Chocolate bar
Phrase 11	them all
Phrase 12	Spose to
Phrase 13	try it
Phrase 14	~~what about~~ water bottle
Phrase 15	~~Hurry~~ animal
Phrase 16	put it away
Phrase 17	mountain air
Phrase 18	~~trader~~ glitter
Phrase 19	a little noise
Phrase 20	To understand
Phrase 21	Their shoes
Phrase 22	I ~~can't keep~~ it
Phrase 23	So it is happen
Phrase 24	~~Snapper~~ there are some apples
Phrase 25	I would have been
Phrase 26	should not ~~Count~~ have gone
Phrase 27	should keep working
Phrase 28	Seher is coming
Phrase 29	You are coming ~~at time~~ aren't you
Phrase 30	gotta be there

FIGURE G.2: Participant EG10 Post-Test

Pre Test: Experiment Group **Student Code** EG12

Instructions:

1- **An audio containing 30 phrases shall be played. Each phrase is made up of two to three words, and it will be repeated thrice.**
2- **Listen carefully and write what you think you heard.**
3- **Don't worry about your answers being wrong. Just write what you think you heard, even if it sounds like gibberish.**

Phrase 1	Talk spark
Phrase 2	input
Phrase 3	~~Teather~~ Teacher
Phrase 4	Ten boys
Phrase 5	The sight
Phrase 6	Have to go
Phrase 7	Have stay
Phrase 8	Good girl
Phrase 9	Next stage
Phrase 10	Night dream
Phrase 11	kiss
Phrase 12	First time
Phrase 13	Less Night
Phrase 14	one
Phrase 15	Hi
Phrase 16	Get out
Phrase 17	Bird
Phrase 18	Later
Phrase 19	Cat
Phrase 20	Just go
Phrase 21	Tell him
Phrase 22	Up and down
Phrase 23	Caliber
Phrase 24	Put back
Phrase 25	~~Put~~ there
Phrase 26	Can go
Phrase 27	unhappy
Phrase 28	~~Bird~~ John says
Phrase 29	is empty
Phrase 30	Convey

FIGURE G.3: Participant EG12 Pre-Test

Post Test: Experiment Group **Student Code** EG12

Instructions:

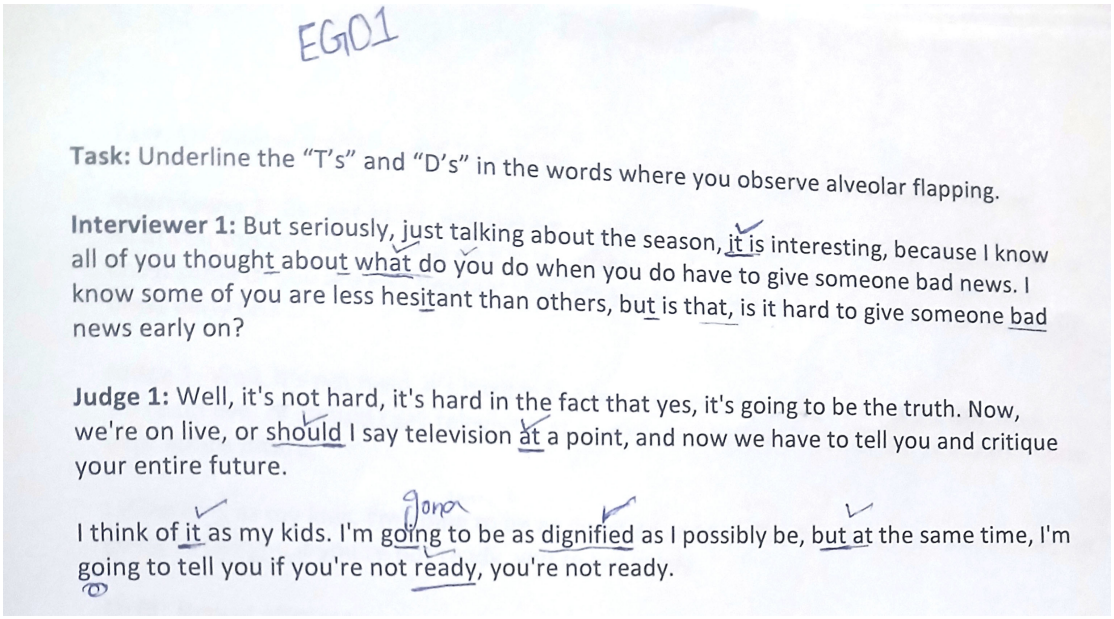
1- **An audio containing 30 phrases shall be played. Each phrase is made up of two to three words, and it will be repeated thrice.**
2- **Listen carefully and write what you think you heard.**
3- **Don't worry about your answers being wrong. Just write what you think you heard, even if it sounds like gibberish.**

Phrase 1	Don't ~~spining~~ Dog's fighting
Phrase 2	in case
Phrase 3	Would you
Phrase 4	Template
Phrase 5	delight
Phrase 6	Toxic
Phrase 7	Bringing
Phrase 8	This year
Phrase 9	next week
Phrase 10	Choclates bar
Phrase 11	~~Them~~ add
Phrase 12	~~Specifier~~ spoose too
Phrase 13	~~Courage~~ the orange
Phrase 14	water ~~bottle~~
Phrase 15	Party animal
Phrase 16	Put it away
Phrase 17	~~Enemy~~ Mountain air
Phrase 18	~~Clever~~ later
Phrase 19	a little noise
Phrase 20	from the start
Phrase 21	their shoes
Phrase 22	I Can't help it
Phrase 23	~~folder~~ ~~paper~~ salt and Pepper
Phrase 24	~~paper~~ There are some apples
Phrase 25	out of ~~them~~ I would have been
Phrase 26	should'nt have gone
Phrase 27	she's been working
Phrase 28	Sehar is coming
Phrase 29	you're coming ~~i check~~ aren't you
Phrase 30	Gotta ~~car~~ be there

FIGURE G.4: Participant EG12 Post-Test

Appendix H

Sample Filled Learner Worksheets


EG01

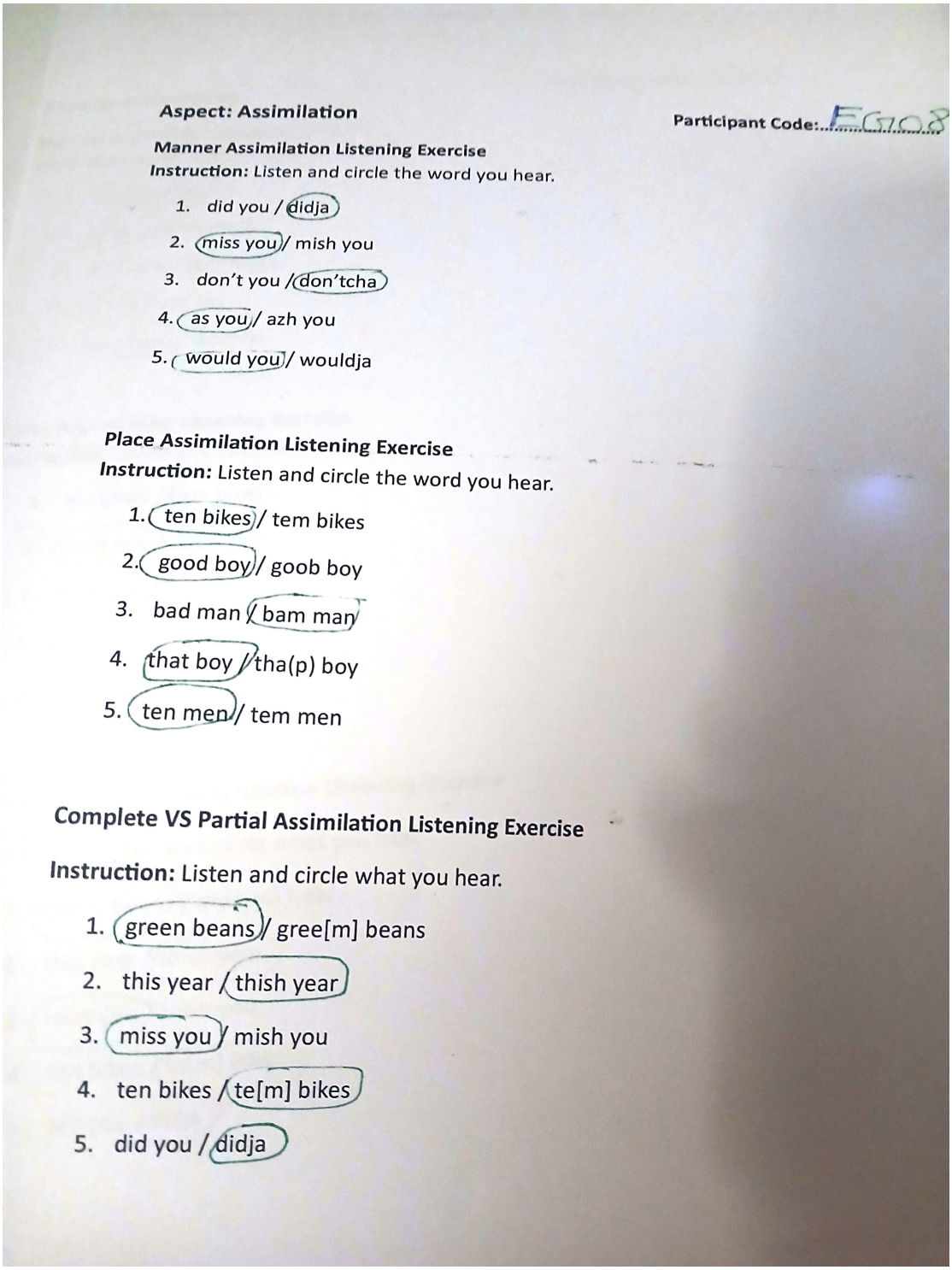
Task: Underline the "T's" and "D's" in the words where you observe alveolar flapping.

Interviewer 1: But seriously, just talking about the season, it is interesting, because I know all of you thought about what do you do when you do have to give someone bad news. I know some of you are less hesitant than others, but is that, is it hard to give someone bad news early on?

Judge 1: Well, it's not hard, it's hard in the fact that yes, it's going to be the truth. Now, we're on live, or should I say television at a point, and now we have to tell you and critique your entire future.

I think of it as my kids. I'm going to be as dignified as I possibly be, but at the same time, I'm going to tell you if you're not ready, you're not ready.

FIGURE H.1: This demonstrates the student's understanding of the alveolar Flapping



Aspect: Assimilation Participant Code: EG08

Manner Assimilation Listening Exercise

Instruction: Listen and circle the word you hear.

1. did you / didja
2. miss you / mish you
3. don't you / don'tcha
4. as you / azh you
5. would you / wouldja

Place Assimilation Listening Exercise

Instruction: Listen and circle the word you hear.

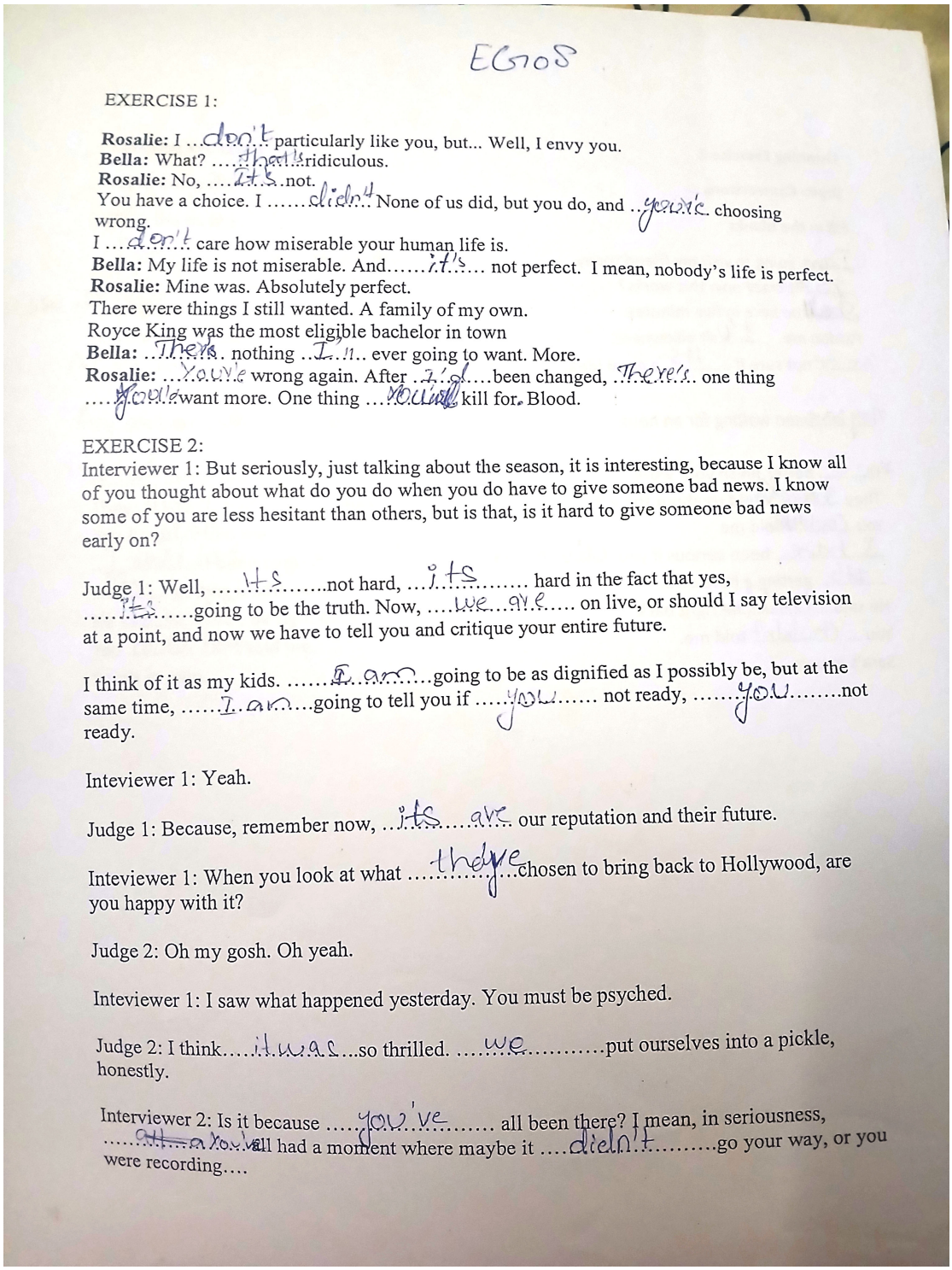
1. ten bikes / tem bikes
2. good boy / goob boy
3. bad man / bam man
4. that boy / tha(p) boy
5. ten men / tem men

Complete VS Partial Assimilation Listening Exercise

Instruction: Listen and circle what you hear.

1. green beans / gree[m] beans
2. this year / thish year
3. miss you / mish you
4. ten bikes / te[m] bikes
5. did you / didja

FIGURE H.2: Student identifying the correct form after listening to the audio



EGOS

EXERCISE 1:

Rosalie: I …don't… particularly like you, but... Well, I envy you.
Bella: What? ….that's…ridiculous.
Rosalie: No, ….it's….not.
You have a choice. I ……didn't…. None of us did, but you do, and ..you're.. choosing wrong.
I …don't… care how miserable your human life is.
Bella: My life is not miserable. And……it's…… not perfect. I mean, nobody's life is perfect.
Rosalie: Mine was. Absolutely perfect.
There were things I still wanted. A family of my own.
Royce King was the most eligible bachelor in town
Bella: ..There's.. nothing ..I'm.. ever going to want. More.
Rosalie: …You've wrong again. After ..I'd….been changed, ..There's.. one thing ….You'll…want more. One thing …you will kill for. Blood.

EXERCISE 2:
Interviewer 1: But seriously, just talking about the season, it is interesting, because I know all of you thought about what do you do when you do have to give someone bad news. I know some of you are less hesitant than others, but is that, is it hard to give someone bad news early on?

Judge 1: Well, ……it's……..not hard, …it's…… hard in the fact that yes, ……it's…..going to be the truth. Now, ….we're….. on live, or should I say television at a point, and now we have to tell you and critique your entire future.

I think of it as my kids. ……I am…going to be as dignified as I possibly be, but at the same time, ……I am…going to tell you if …..you…… not ready, ……you……not ready.

Inteviewer 1: Yeah.

Judge 1: Because, remember now, ..it's…..are our reputation and their future.

Inteviewer 1: When you look at what ……they've…chosen to bring back to Hollywood, are you happy with it?

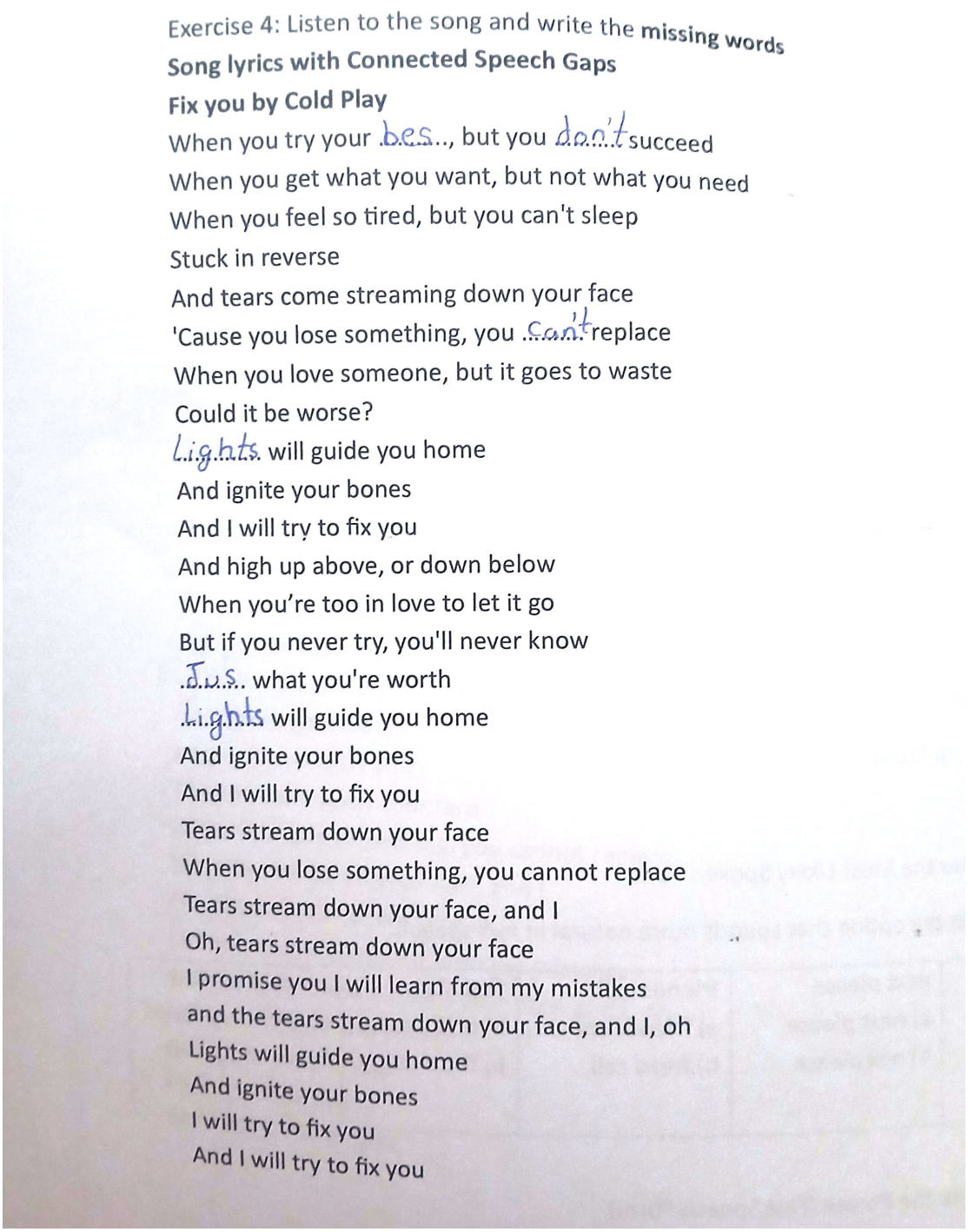
Judge 2: Oh my gosh. Oh yeah.

Inteviewer 1: I saw what happened yesterday. You must be psyched.

Judge 2: I think…..it was…so thrilled. ……we……put ourselves into a pickle, honestly.

Interviewer 2: Is it because ……you've…… all been there? I mean, in seriousness, ……all…you've all had a moment where maybe it ….didn't……go your way, or you were recording….

FIGURE H.3: Student listens to the audio and writes down the contractions



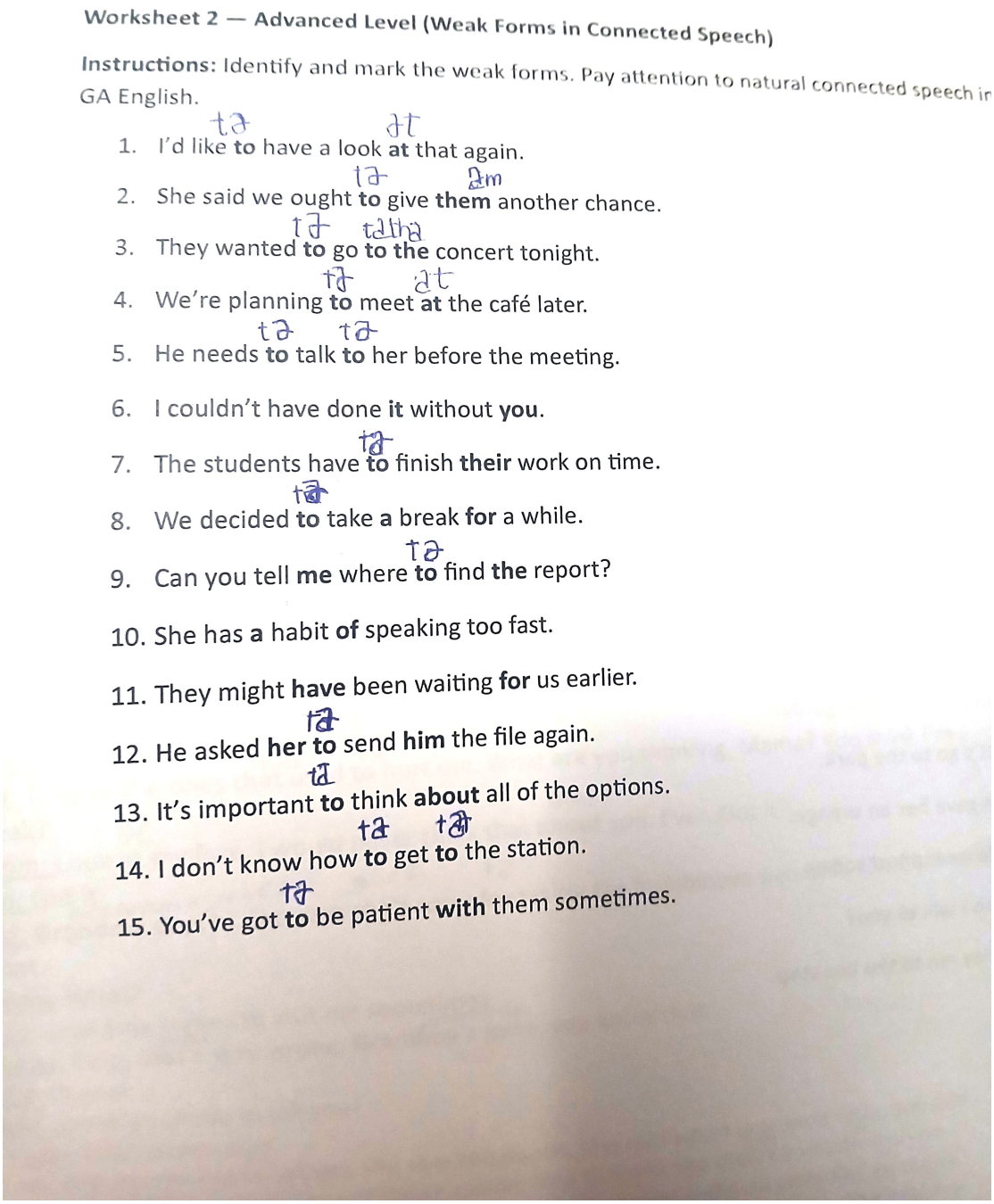
Exercise 4: Listen to the song and write the **missing words**

Song lyrics with Connected Speech Gaps

Fix you by Cold Play

When you try your bes.., but you don't succeed
When you get what you want, but not what you need
When you feel so tired, but you can't sleep
Stuck in reverse
And tears come streaming down your face
'Cause you lose something, you can't replace
When you love someone, but it goes to waste
Could it be worse?
Lights will guide you home
And ignite your bones
And I will try to fix you
And high up above, or down below
When you're too in love to let it go
But if you never try, you'll never know
Just what you're worth
Lights will guide you home
And ignite your bones
And I will try to fix you
Tears stream down your face
When you lose something, you cannot replace
Tears stream down your face, and I
Oh, tears stream down your face
I promise you I will learn from my mistakes
and the tears stream down your face, and I, oh
Lights will guide you home
And ignite your bones
I will try to fix you
And I will try to fix you

FIGURE H.4: Student's responses shows how they perceive the elided forms



Worksheet 2 — Advanced Level (Weak Forms in Connected Speech)

Instructions: Identify and mark the weak forms. Pay attention to natural connected speech in GA English.

1. I'd like **to** have a look **at** that again.
2. She said we ought **to** give **them** another chance.
3. They wanted **to** go **to the** concert tonight.
4. We're planning **to** meet **at** the café later.
5. He needs **to** talk **to** her before the meeting.
6. I couldn't have done **it** without **you**.
7. The students have **to** finish **their** work on time.
8. We decided **to** take **a** break **for** a while.
9. Can you tell **me** where **to** find **the** report?
10. She has **a** habit **of** speaking too fast.
11. They might **have** been waiting **for** us earlier.
12. He asked **her to** send **him** the file again.
13. It's important **to** think **about** all of the options.
14. I don't know how **to** get **to** the station.
15. You've got **to** be patient **with** them sometimes.

FIGURE H.5: This shows their grasp on the concept of weak forms and the use of schwa

Appendix I

Ethical Approval

To: The Head of Department
Department of English
Abasyn University

Subject: Application for Permission to Conduct Experimentation on Connected Speech

Respected Sir,

I respectfully request permission to conduct experimentation related to my research on *Connected Speech*. The study will focus on selected aspects such as assimilation, elision, contractions, weak forms, and alveolar flapping.

The purpose is to collect data for my research on listening comprehension and cognitive load in ESL learners. I assure you that all activities will be conducted ethically and without disruption to regular classes.

Your approval to proceed with this research phase will be greatly appreciated.

Yours sincerely,
Bushra Sikander
Lecturer, Department of English
Abasyn University

Date:.....

Allowed

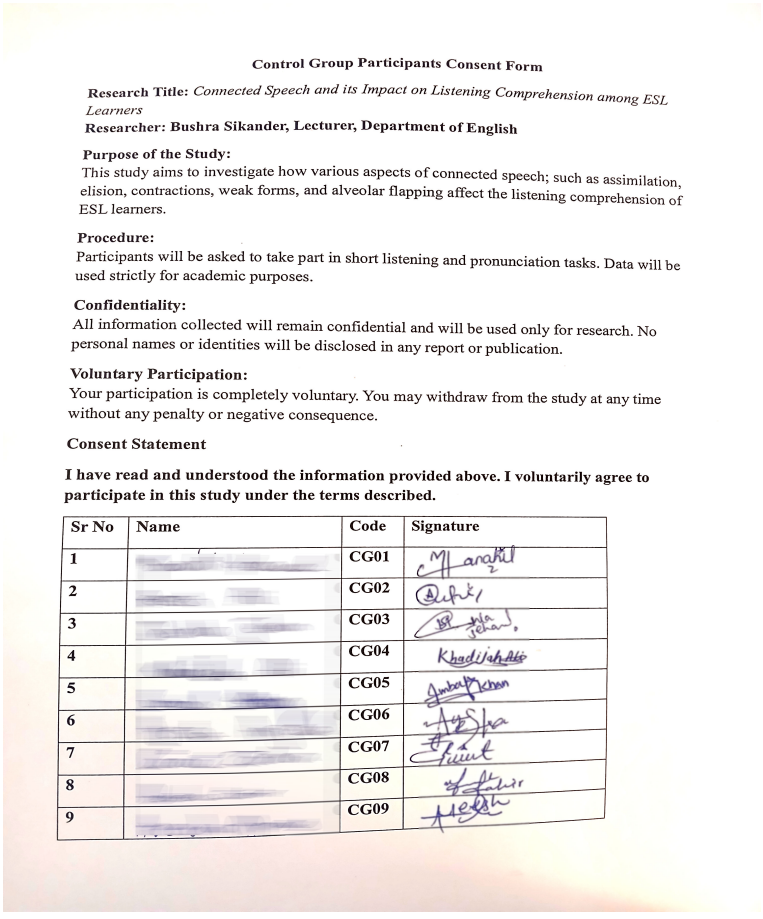
HOD's Name: Dr. Professor Raja Nasim Akhter

HOD's Signature:..... 24/11/25

FIGURE I.1: Consent Form Signed by HOD, English Department, Abasyn University

Appendix J

Student Consent Forms



Control Group Participants Consent Form

Research Title: *Connected Speech and its Impact on Listening Comprehension among ESL Learners*
Researcher: Bushra Sikander, Lecturer, Department of English

Purpose of the Study:
This study aims to investigate how various aspects of connected speech; such as assimilation, elision, contractions, weak forms, and alveolar flapping affect the listening comprehension of ESL learners.

Procedure:
Participants will be asked to take part in short listening and pronunciation tasks. Data will be used strictly for academic purposes.

Confidentiality:
All information collected will remain confidential and will be used only for research. No personal names or identities will be disclosed in any report or publication.

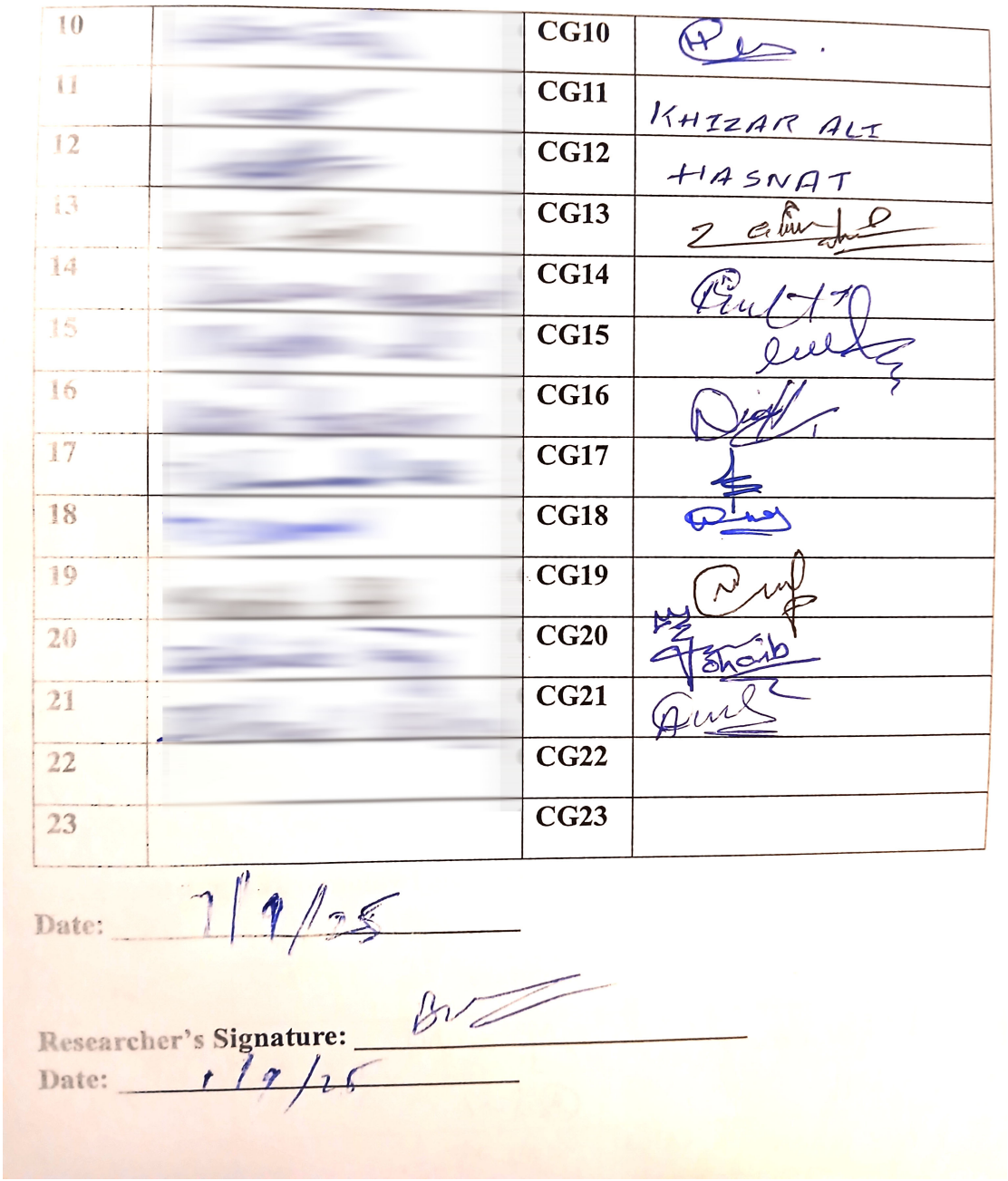
Voluntary Participation:
Your participation is completely voluntary. You may withdraw from the study at any time without any penalty or negative consequence.

Consent Statement

I have read and understood the information provided above. I voluntarily agree to participate in this study under the terms described.

Sr No	Name	Code	Signature
1		CG01	
2		CG02	
3		CG03	
4		CG04	
5		CG05	
6		CG06	
7		CG07	
8		CG08	
9		CG09	

FIGURE J.1: Conrol-Group Participants' Consent Terms and Form



10	[redacted]	CG10	
11	[redacted]	CG11	KHIZAR ALI
12	[redacted]	CG12	HASNAT
13	[redacted]	CG13	
14	[redacted]	CG14	
15	[redacted]	CG15	
16	[redacted]	CG16	
17	[redacted]	CG17	
18	[redacted]	CG18	
19	[redacted]	CG19	
20	[redacted]	CG20	Shoaib
21	[redacted]	CG21	
22		CG22	
23		CG23	

Date: 7/9/25

Researcher's Signature: _____

Date: 7/9/25

FIGURE J.2: Conrol-Group Participants' Consent Form

Experiment Group Participant Consent Form

Research Title: *Connected Speech and its Impact on Listening Comprehension among ESL Learners*

Researcher: *Bushra Sikander, Lecturer, Department of English*

Purpose of the Study:
This study aims to investigate how various aspects of connected speech affect the listening comprehension of ESL learners.

Procedure:
Participants will be asked to take part in short listening and pronunciation tasks. Data will be used strictly for academic purposes.

Confidentiality:
All information collected will remain confidential and will be used only for research. No personal names or identities will be disclosed in any report or publication.

Use of Voice Recorder

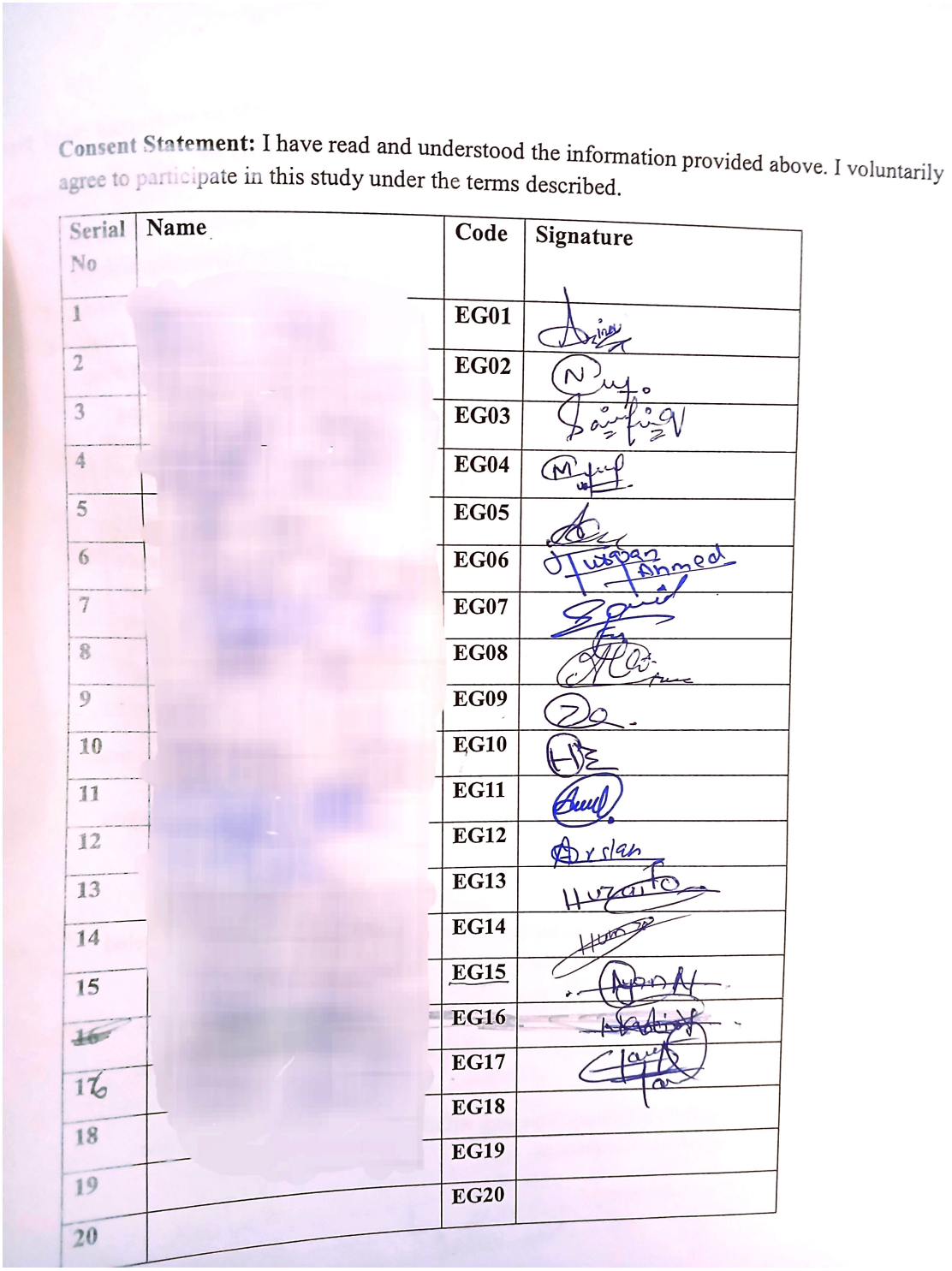
With your permission, we will record your voice during the session to capture verbal responses and interactions. These recordings will be used solely for research purposes, including transcription, analysis, and reporting. A short video will be recorded from behind to keep you anonymous.

Confidentiality

- Your identity will remain confidential. Audio files will be stored securely and labeled with a participant code rather than your name.
- Only authorized members of the research team will have access to the recordings.
- Recordings will not be shared publicly or used in any way that could identify you.

Voluntary Participation:
Your participation is completely voluntary. You may withdraw from the study at any time without any penalty or negative consequences.

FIGURE J.3: Experiment-Group participants' Consent Terms



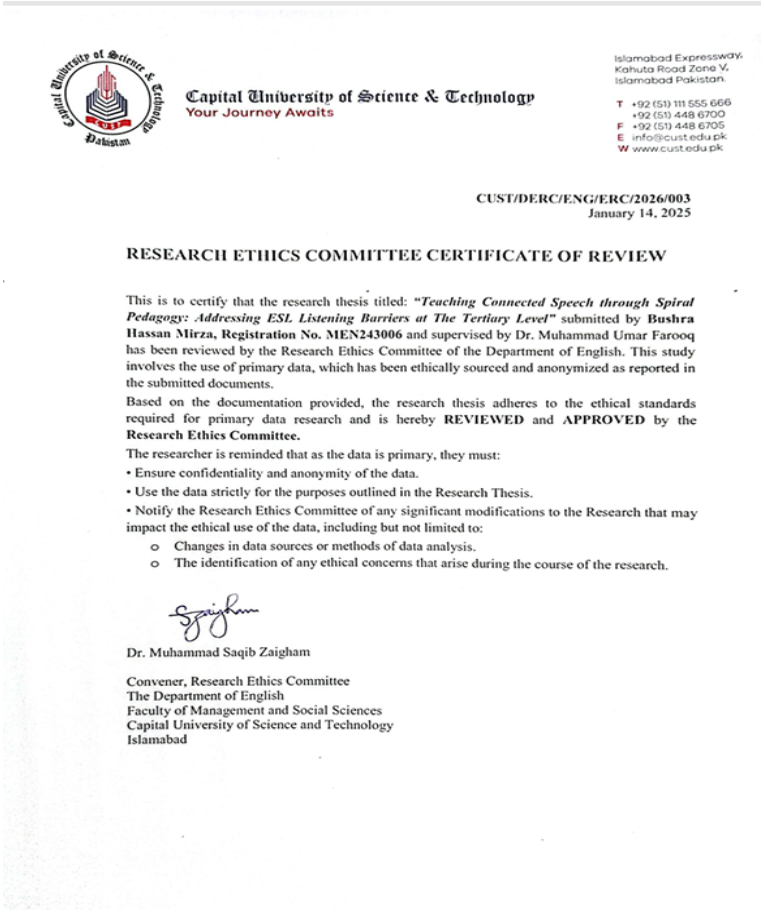
Consent Statement: I have read and understood the information provided above. I voluntarily agree to participate in this study under the terms described.

Serial No	Name	Code	Signature
1		EG01	
2		EG02	
3		EG03	
4		EG04	
5		EG05	
6		EG06	
7		EG07	
8		EG08	
9		EG09	
10		EG10	
11		EG11	
12		EG12	
13		EG13	
14		EG14	
15		EG15	
~~16~~		EG16	
17		EG17	
18		EG18	
19		EG19	
20		EG20	

FIGURE J.4: Experiment-Group participants' Consent Form

Appendix K

Certificate of Review



Capital University of Science & Technology
Your Journey Awaits

Islamabad Expressway,
Kahuta Road Zone V,
Islamabad Pakistan.

T +92 (51) 111 555 666
+92 (51) 448 6700
F +92 (51) 448 6705
E info@cust.edu.pk
W www.cust.edu.pk

CUST/DERC/ENG/ERC/2026/003
January 14, 2025

RESEARCH ETHICS COMMITTEE CERTIFICATE OF REVIEW

This is to certify that the research thesis titled: ***"Teaching Connected Speech through Spiral Pedagogy: Addressing ESL Listening Barriers at The Tertiary Level"*** submitted by **Bushra Hassan Mirza, Registration No. MEN243006** and supervised by Dr. Muhammad Umar Farooq has been reviewed by the Research Ethics Committee of the Department of English. This study involves the use of primary data, which has been ethically sourced and anonymized as reported in the submitted documents.

Based on the documentation provided, the research thesis adheres to the ethical standards required for primary data research and is hereby **REVIEWED** and **APPROVED** by the **Research Ethics Committee.**

The researcher is reminded that as the data is primary, they must:

- Ensure confidentiality and anonymity of the data.
- Use the data strictly for the purposes outlined in the Research Thesis.
- Notify the Research Ethics Committee of any significant modifications to the Research that may impact the ethical use of the data, including but not limited to:
  - Changes in data sources or methods of data analysis.
  - The identification of any ethical concerns that arise during the course of the research.

Dr. Muhammad Saqib Zaigham

Convener, Research Ethics Committee
The Department of English
Faculty of Management and Social Sciences
Capital University of Science and Technology
Islamabad

FIGURE K.1: Research Ethics Committee Certificate of Review