

LOCAL IDENTITY, GLOBAL VOICE: IMPROVING SPEAKING SKILLS OF ESL LEARNERS THROUGH THE CULTURALLY RESPONSIVE-AUDIO DESCRIPTION (CRAD) MODEL

¹*Dr. Farah Zaib** and ²*Dr. Shahid Nawaz*

¹*Lecturer, Department of English Linguistics, The Islamia University of Bahawalpur*

²*Assistant Professor, Department of English Linguistics, The Islamia University of Bahawalpur*

**Corresponding Author: 2050zaib.un.nisa@gmail.com*

Abstract

This mixed-methods study investigated the efficacy of the Culturally Responsive Teaching- Audio Description (CRAD) model in improving the speaking skills of undergraduate ESL learners. Grounded in the Affective Filter Hypothesis (Krashen, 1985) and Culturally Responsive Teaching (CRT), the research hypothesized that leveraging local Pakistani heritage content would reduce speaking anxiety and provide a linguistic scaffold for complex oral production. The quantitative (experimental) phase (N=60) used a pre-test/post-test design, measuring fluency, linguistic complexity (T-unit analysis), and coherence. Results demonstrated statistically significant improvements for the CRAD group ($p < .05$) across all metrics, particularly in descriptive vocabulary and complex syntax. The subsequent qualitative phase analyzed student reflections (N=20), yielding themes of Reduced Speaking Inhibition (lowered anxiety), Targeted Linguistic Precision (necessitated use of complex structures), and Motivation and Refinement. The findings affirm that CRAD effectively bridges affective safety and linguistic complexity. The study concludes that leveraging students' local identity transforms the ESL speaking task into an empowering communicative act, offering a data-backed pedagogical model for advanced oral skill enhancement.

Keywords: *Culturally Responsive Teaching (CRT), Audio Description (AD), ESL Speaking Skills, Local Identity, Affective Filter, Linguistic Complexity.*

Background of the study

Despite continued pedagogical efforts, persistent deficiencies in the oral communicative competence and confidence of undergraduate ESL learners indicate the insufficiency of conventional teaching approaches. Consequently, innovative pedagogical approaches are imperative (Krashen, 1985; Ladson-Billings, 1995). This research addresses a significant gap in English as a second language (ESL) curriculum by advocating a shift from mere linguistic correction to the cultivation of confident and globally-aware communication. Through the validation of integrated Culturally Responsive Teaching-Audio Description (CRAD) model, this research simultaneously aims to enhance the objective speaking performance while nurturing a confident "Global Voice." The integrated model is essential for providing learners with the cultural security needed to help reduce affective filters (Krashen, 1985), and the structured scaffolding necessary to foster precise and complex language production (Vygotsky, 1978). The study's variables are discussed below in this context.

Culturally Responsive Teaching (CRT)

Culturally responsive teaching (CRT) forms a school of educational thought and practice in which the diverse cultural backgrounds, experiences, and perspectives of learners are recognized and valued, and their multi-dimensional aspects are interwoven throughout every aspect of pedagogy to create inclusive, equitable, and engaging learning environments. CRT goes beyond the generic "good teaching" by consciously incorporating students' cultural identities and lived experiences into the construction of curricula, the teaching and learning processes, and interactions in the classroom in ways that promote academic achievement, foster

positive self-identity, and nurture critical thinking and social justice (Caingcoy, 2023; Gay, 2002; Ladson-Billings, 2021; Smith et al., 2022). Key principles of CRT are based on interconnected core elements that are necessary for effective application. Central among these is the Recognition and Integration of Cultural Backgrounds (Caingcoy, 2023; Gay, 2002) which intentionally weaves students heritage, experience, and perspective into the learning continuum. This is very much in line with the Building of Relationships and Community (Bennett et al., 2018; Gunn et al., 2020), which is the development of an inclusive and supportive environment where learners feel valued. Effective CRT further requires Critical Reflection and Teacher Self-awareness (Ladson-Billings, 2021; Smith et al., 2022), which requires educators to examine their own biases and instructional practices. Teachers are required to engage in Adapting Instruction and Curriculum (Gay, 2002; Mehta, 2024), where materials and pedagogies are altered in order to make content culturally relevant and accessible. Furthermore, CRT seeks to Promote Critical Consciousness (Caingcoy, 2023; Wesley-Nero, 2024), which encourages learners to analyze and challenge social inequities. Sustained efficacy is maintained through ongoing Professional Development (Morrison et al., 2021; Warren, 2018) thus having educators constantly refine their culturally responsive competencies by using multicultural literature, engaging students' families, creating spaces for culturally reflective learning, and designing lessons that link academic content to students' real-world encounters (Bennett et al., 2018; Gunn et al., 2020; Nawaz et al., 2021a,b).

Audio Description (AD)

Audio description (AD) is a narrative technique intended to translate visual media into verbal form by systematically defining the who, what, where, and how of events depicted in a visual media in order to make visual media accessible to all audiences (Maszerowska et al., 2014). Over the past two decades, the pedagogical potential of AD in the context of translation studies and language acquisition has been identified by an ever-growing number of scholars. AD has been described as a pedagogical approach to language teaching (Rodrigues-Barbosa, 2013) and has been characterized as a novelty to increase linguistic competencies (Burger, 2016). Herrero and Escobar (2018) proposed incorporating audio description and film literacy instruction into the language teaching curriculum. They reiterated the significance of audio description and contended that audio description initiatives might enhance language learners' intercultural, cultural, and linguistic proficiencies (Nawaz et al., 2022).

Task-based Language Teaching

The term "task" has different connotations for different people (Leaver & Willis, 2004), but most people would agree that the task must accomplish a certain goal or culminate in some conclusion. Nunan (2003) claims that a task is a work done in class in which students manipulate, grab, create, or interact with a foreign language while focusing more on the meaning than the form. Willis and Willis (2007) defines tasks as authentic goal-oriented actions that result in an outcome. In task-based language teaching (TBLT), learners use a second language (L2) to complete their communicative tasks through a task or activity. Emerging from the communicative approach in the 1980s, TBLT is a learner-centered teaching-learning approach (Richards & Rogers, 2014). The teacher engages the learners in meaningful tasks in which they are able to use authentic L2 and learn to perform functions in real life, such as asking for directions, requesting information, accepting an invitation or rejecting an invitation, writing official letters, reading maps. In TBLT the main purpose is to engage the learners in meaningful tasks that engage them use language in real life communication. These activities are not random and are sequenced from easy to challenging levels, depending on the capabilities of the learners.

Speaking Skills

Speaking is a very engaging and productive skill and plays a significant role in our lives. We

communicate our ideas, feelings, and thoughts and interact with others through speech. A speaker's ability is often measured in terms of their fluency and accuracy. The speaking skill is often used to test overall communication skills. Savignon (1991) postulates that the act of speaking is evidence of an individual's degree of competence, which she defines as the ability to use language for interactive purposes as opposed to the ability to perform discrete-point grammar tests. We spend 30% of our daily activities on verbal communication (Palmer, 2014). In order to be an effective communicator, he should be well-versed in four aspects of speaking skills: phonology, speech functions, interactional skill and discourse skills. Speaking especially becomes difficult when one is forced to speak in a foreign or second language. Speaking skills include four areas: linguistic competence, strategic competence, communicative competence, and social competence (Littlewood, 1981). Speaking is not simply an imitation of sounds or verbalization of written text (Bygate, 1987). Effective performance of this productive skill requires utilization of both compensative and facilitative techniques. When it comes to verbal communication, the learner has to know a few things, such as how to resolve issues resulting from communication and how to negotiate and control interaction. Proficiency in verbal communication is only learned through actual experience.

Problem Statement

The persistent difficulty in advanced English as a Second Language education is the simultaneous lack of communicative speaking skills (limited fluency, limited linguistic complexity and limited coherence) and salient affective impediments (high anxiety and low speaking confidence) among the undergraduate ESL learners. Conventional pedagogical approaches often do not address this dual dilemma because they do not systematically use the personal and local cultural resources of students to make communication salient and reassuring. Consequently, there is a critical pedagogical and empirical gap regarding the validation of an integrated, asset-based strategy—the Culturally Responsive Teaching-Audio Description (CRAD) model—that can demonstrably and simultaneously augment objective speaking performance and foster the self-assured "Global Voice" requisite for academic and professional success. The following research objectives stem from this research problem.

Research Objectives

1. To statistically determine the effect of the Culturally Responsive Teaching-Audio Description (CRAD) model on improving the speaking skills (fluency, linguistic complexity, and coherence) of undergraduate ESL learners compared to a conventional teaching method.
2. To explore and analyze undergraduate ESL learners' reflections to understand how the integration of local cultural content influences their speaking confidence, anxiety levels, and motivation to refine their speaking skills ("Global Voice").

Hypotheses

H1: Undergraduate ESL learners taught using the Culturally Responsive Teaching-Audio Description (CRAD) model will show a statistically significant improvement in their speaking skills (fluency, linguistic complexity, and coherence) compared to those taught using a conventional teaching method.

H0: There is no statistically significant difference in the speaking skills (fluency, linguistic complexity, and coherence) of undergraduate ESL learners taught using the Culturally Responsive Teaching-Audio Description (CRAD) model and those taught using a conventional teaching method.

The Significance of the Research

The importance of this study lies in its pioneering contribution to pedagogy by elaborating and empirically testing the novel CRAD model as an advanced, high-impact instructional model for undergraduate ESL speaking curricula. It also deals with affective barriers by showing that the level of speaking anxiety can be significantly reduced by including culturally relevant

content, while at the same time increasing the confidence and motivation of the learners. This shift in focus moves beyond linguistic correction to the imperative of authentic communication. Furthermore, the study provides statistical evidence of the impact of CRAD on objective speaking proficiencies (i.e. fluency, complexity, and coherence), which fills in a significant gap related to the concomitant use of CRT and Audio Description (AD). Ultimately, these findings are instrumental to practice, providing information for language educators, curriculum design, and teacher training programs on how best to achieve efficacious integration of local culture in order to develop the advanced "Global Voice" in ESL learners.

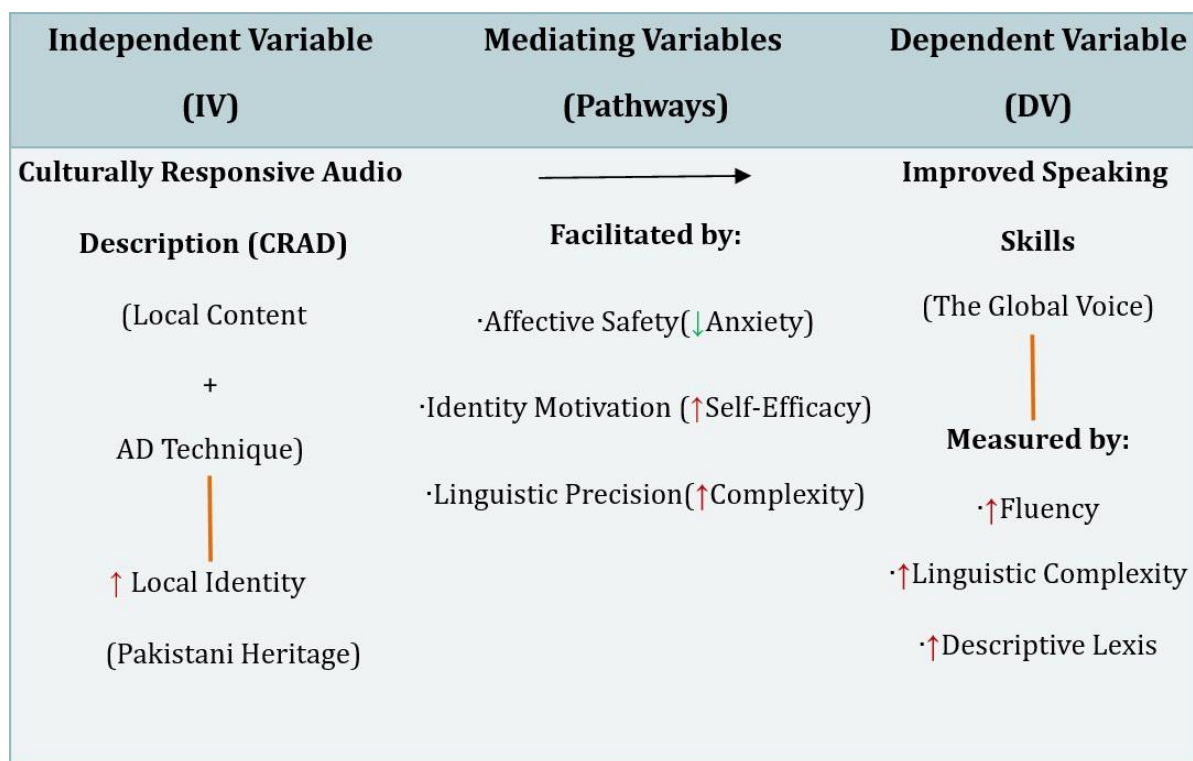
Theoretical Framework

The present study is based on an integrated theoretical framework that deals with the cognitive, affective, and social aspects of advanced speaking skill acquisition. Central to this framework is the Affective Filter Hypothesis (Krashen, 1985), which posits that high affective factors can prevent language input from reaching the language acquisition device. The hypothesis explains how the use of Local Identity can reduce anxiety and inhibition, thereby facilitating the intake and output of language. Complementing this affective security is Deci & Ryan's Self - Determination Theory (1985), which acts as a motivational construct. This theory demonstrates how perceived cultural proficiency further boosts intrinsic motivation and self-efficacy, thus galvanizing efforts to fine-tune speaking skills. Finally, the necessary cognitive structure is provided by the synergy of Vygotsky's Sociocultural Theory (1978), and the principles of Task-Based Language Teaching (Willis & Willis, 2007), which together form a linguistic scaffold. This scaffold details how the structured demands of Audio Description challenge learners to move beyond simple conversational exchanges and produce the complex, precise linguistic output necessary for the "Global Voice."

Conceptual Framework

The research uses a linear model to evaluate the effectiveness of the Culturally Responsive-Audio Description (CRAD) model. The independent variable is the CRAD model itself that is defined as the synergistic amalgamation of the Local Content, specifically the Local Identity derived from Pakistani heritage, within structured Audio Description task. Local Identity is the main impetus in the intervention. The path from intervention to ultimate goal is described by three interacting mediating variables: Affective Safety, in which local content reduces anxiety and lowers the affective filter level; Identity Motivation, in which cultural expertise increases the levels of intrinsic motivation and self-efficacy for authoritative discourse; and Linguistic Precision, in which the structured AD task increases linguistic complexity. All these pathways together culminate in the dependent variable—Improved Speaking Skills—conceptually defined as "The Global Voice." Improvement is concretely evaluated by the three linguistic metrics: fluency, linguistic complexity and descriptive lexis. In other words, the framework suggests that the CRAD model uses Local Identity to have a positive effect on the affective states and motivation of the learners, thus giving rise to the production of accurate language and measurable improvements in speaking skills.

Figure 1.
CRAD Conceptual Framework (Zaib, 2025)



Literature Review

CRT and ESL Speaking Skills

The integration of culture into language instruction has historically been associated with positive results in communicative and engagement proficiency. For example, Menacho-Vargas et al. (2021), focusing on Peruvian English as a foreign language (EFL) learners, found that culturally based dialogues contributed to significant development in speaking skills when compared to traditional and non-contextualized instruction. This finding underscores the pedagogical relevance of exploiting the learners' background. Recent studies have focused on CRT's ability to promote motivational outcomes, especially in the context of higher education in Asian countries. Idrus et al. (2023) explored preservice teachers' experiences with CRT in Malaysian ESL settings, concluded that the approach led to more fluent and active speakers, noting that students became less shy and more willing to express opinions. Jalil (2023), working in Malaysian English classrooms, found that Culturally Responsive Language Pedagogy increased motivation and engagement, consequently enhancing speaking skills and bridging cultural gaps. Reskyani et al. (2024), who conducted research in the higher education environment in Indonesia, showed that the use of a CRT approach was created a positive environment in class, directly increasing motivation and interest in speaking practice. Likewise, Matiso (2024), focusing on CRT optimization in South African multicultural ESL classrooms, identified inclusive, culturally responsive strategies (i.e. peer collaboration and the incorporation of technology) that proved effective in achieving higher engagement and improved general oral participation. Collectively, these studies support the idea that the inclusion of cultural relevance into pedagogy serves as a powerful motivational force, leading to lower affective barriers and demonstrable improvements in fluency, active engagement, and confidence.

CRT and Speaking Skills in Pakistani Context

A salient theme in Pakistani literature is the benefit of leverage the knowledge of the language and culture already known to learners in order to develop confidence. Amjad et al. (2024), drawing on bilingual and CRT approaches, concluded that strategic use of the L1 has a great impact on enhancing student confidence and fluency in L2 production. Supporting such an understanding of engagement, Rafi and Fox (2021), researched translanguaging as a form of CRT in university settings and reported that the active promotion of engagement through aiding learners to use the mother tongue in learning, is central to the development of more spontaneous speaking practices. Further, Nisar et al. (2023) emphasized the critical need for cultural sensitivity in ESL classrooms, requiring instruction to align with the local socio-cultural norms to ensure efficacy.

Beyond confidence, other researches also support the efficacy of contextual and interactive methods for improving specific speaking skills. Studies by Bango et al. (2023), Nosheeen et al. (2020), Samiullah (2018), and Yasmeen et al. (2023) collectively affirm the effectiveness of contextual and interactive approaches in addressing communication skills deficiencies.

The extant literature has affirmed the importance of Culturally Responsive Teaching in producing positive affective and general engagement outcomes; nevertheless, most of the studies focus on general speaking strategies or on L1/translanguaging use. A critical gap exists because no reviewed studies (especially in the Pakistani context) specifically worked on the rigorous and structured demands of the Audio Description technique when amalgamated to the rigors of CRT to target quantifiable linguistic outcomes such as, linguistic complexity (quantified by T-unit analysis) and coherence. This gap validates the need for the present study to affirm the Culturally Responsive-Audio Description model as a specific strategy to move local confidence to precise, globally-relevant communication. The research gaps that have been identified are multifaceted: from a theoretical perspective, no prior research connects the Affective Filter Hypothesis and CRT with the Audio Description technique to build the CRAD pedagogical model. Methodologically speaking, no previous work has been conducted with the same structured, task-based methodology, which makes it different from general conversational practice. From a contextual perspective, the efficacy of the CRAD model has not yet been evaluated in the challenging context of English as a Second Language (ESL) learners in undergraduate education where the level of linguistic complexity and self-confidence is the key to deciding suitability to continue in a professional capacity. Scope-wise, although existing research focuses on engagement in general, this study focuses on both affective domains (confidence and anxiety) and objective productive skills (fluency, linguistic complexity, and coherence) simultaneously and systematically. Finally, the scope of the investigation resolved the depth gap by going beyond single methodology validation towards a comprehensive, rigorous, mixed methodology investigation, thus offering a fine-grained linguistic and affective evidence for the CRAD model.

Related Studies on AD and Speaking Skills

Ibáñez-Moreno and Vermeulen (2013, 2014, 2015a, 2015b, 2015c) have carried out a series of investigations on the use of Audio Description (AD) as a pedagogical instrument for language acquisition. The first experiment conducted with Dutch learners at the B2 level of Spanish learning in 2013 proved that AD had a significant impact in improving lexical and phraseological competence, and it also enhanced the learner's motivation. Subsequently, a case study was conducted in 2014 in which they investigated the influence of AD on integrated linguistic skills, i.e. listening, reading, speaking and writing skills, and intercultural competence in a direct, in-person context. Building on these results, the researchers created the VIDEOS for Speaking (VISp) mobile application that was originally used in a study (Ibáñez-Moreno & Vermeulen, 2015a) to test the application's potential in teaching English to Spanish

students (these initial results showed that there were positive effects of using the application on motivation, and that there was some benefit in vocabulary learning). Follow-up studies by the same researchers in 2015b and 2015c confirmed the effectiveness of the redesigned VISP app in improving oral production, grammatical accuracy and verbal proficiency in English of Spanish learners, and in turn, confirmed the potential of this app as a valuable and motivating tool for both online and traditional classroom instruction.

Cenni and Izzo (2016) conducted an exploratory study on B2 - level Italian as a foreign language learners, focusing on AD script composition in the target language as opposed to revoicing. The results of their study indicated that script construction was effective for fostering linguistic consciousness and reflection on language. Parallel to this work, Talaván and Lertola (2016) conducted a quasi-experimental study in 2 months with B1 level ESP Tourism students using the ClipFlair software. The results have shown that AD significantly improved integrated language competencies, including the oral production, writing (AD script/forum posts), listening and reading comprehension. Participants also reported an improvement in grammatical accuracy, vocabulary richness, confidence while speaking and sustained motivation, despite challenges such as text condensation.

Navarrete (2018) conducted a small-scale mixed methods experimental study with B1 Spanish learners which also used ClipFlair as a means of interrogation into AD's impact on oral production. Data that was drawn from pre-tests, questionnaires and AD tasks that were evaluated on, for example, fluency and grammaticality, showed that there were positive student perceptions towards the use of AD tasks in comparison to conventional podcasting. These results created a basis for additional exploration of AD within an educational setting.

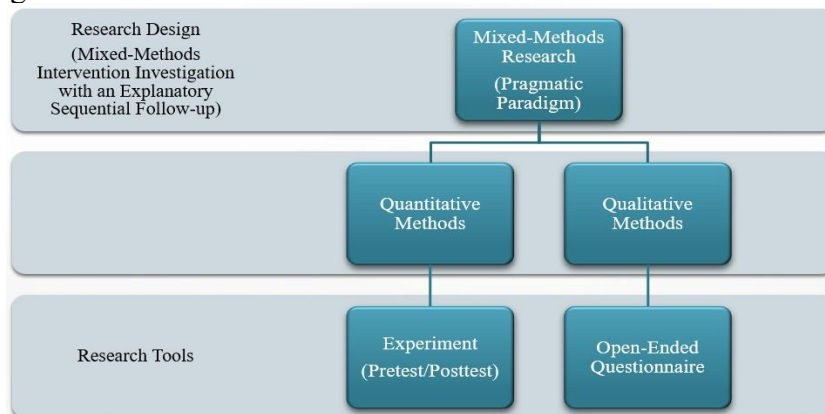
Subsequent studies have both supported and expanded the perceived benefits of AD for language learning. Calduch and Talaván (2018) focused on the writing proficiency of B1-B2 level Spanish learners and found statistically significant improvements in vocabulary retention, stylistic variation, and overall compositional competence due to the use of AD text composing. Concurrently, improvements in oral skills that were not included in the goals were found, especially improved pronunciation and fluency. Ibáñez Moreno and Escobar (2021) showed the success of AD strategies in a collaborative online writing project among ESP tourism students in which video descriptions resulted in better performance on final assessments and motivated students in regard to English writing. Bausells-Espín (2022) also developed the "triple-connection hypothesis" in which the perceived difficulty of AD projects by university students was directly related to the perceived utility and progress in learning in the context of integrated, communicative, and lexical competencies. Finally, Zengin and Yildiz (2024) have also used metaphors analysis to capture the perspectives of students, where it was found that learners linked AD with knowledge creation, creativity, and transfer and thus the emotional and cognitive advantages inherent in this instructional approach within the English language teaching.

Research Methodology

To address the research questions, this study employs a mixed-methods intervention investigation with an explanatory sequential follow-up research design to determine the effects of CRAD on the speaking skills of ESL learners. This methodological design is the most suitable for answering the research questions and, more importantly, for generalizing conclusions from the data obtained. In this mixed-methods study (aligned with the pragmatic paradigm), quantitative and qualitative data are collected through the use of pre-test/post-test and open-ended questionnaire, respectively and analyzed separately before the final integration of the results. The insights gained from the qualitative phase are used to provide a detailed explanation and interpretation of the quantitative findings (Kroll & Neri, 2009).

Figure 2.

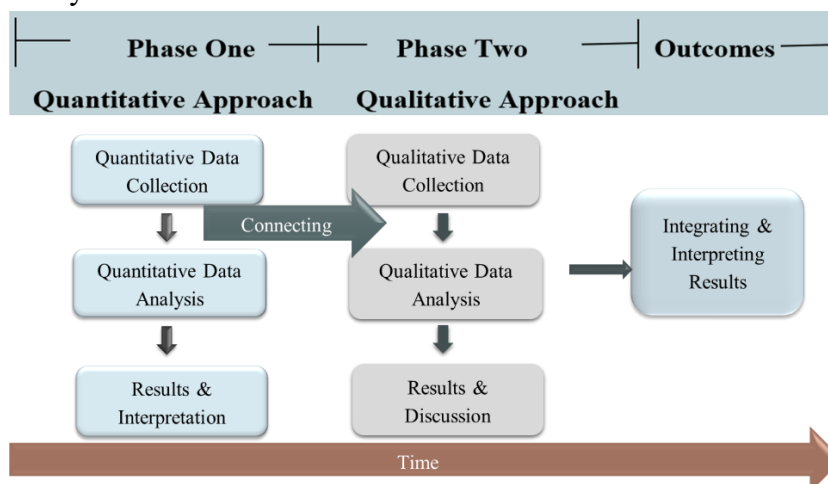
A Mixed-Methods Intervention Investigation with an Explanatory Sequential Follow-Up Research Design



The study was designed in two different phases, Phase One consisted of a quantitative part, followed by a second qualitative phase in which the information obtained was explicitly linked and used to explicate the outcomes of the first phase.

Figure 3.

Phases of the Study



Population, Sample and Sampling

All ESL learners participating in communication skills courses constitute the study's theoretical population. The primary sample for the quantitative phase was then drawn from the accessible population. To ensure that the results are representative and generalizable, this research employs simple random sampling, a type of probability sampling utilized in the experimental phase of the current study. Randomness refers to a specific, scientific procedure in which every individual has an equal chance of being selected (Rubin & Babbie, 2010). The second phase of the study applies purposive sampling, as it allowed for the selection of participants with varied opinions and experiences concerning the phenomenon under investigation (Dörnyei, 2007). A key characteristic of a sample is its representativeness, which is often considered more critical than its size (Ary et al., 2009). The main benefit of using a sample is it makes research more feasible and cost-effective than surveying the entire population. The sample size for the quantitative phase is 72 ESL learners who were enrolled in English Linguistics Department at The Islamia University of Bahawalpur. Purposive sampling was adopted for the study's second phase because it allows the selection of information-rich cases to be studied in depth (Patton, 1990) and provides varied insights into the phenomenon under inquiry (Dörnyei, 2007). The

qualitative phase consisted of 18 ESL learners. The homogeneity of both groups was also established on the basis of the groups size, prior grade marks as well as the current level of proficiency in speaking skills.

Data Collection Tools and Phases

The data were collected in two consecutive phases (i.e., quantitative and qualitative) in accordance with the nature of the explanatory sequential mixed-methods follow-up design (Green & Caracelli, 1997; Creswell, 2003). Therefore, different tools were required for the investigation.

1. TOEIC test for speaking proficiency

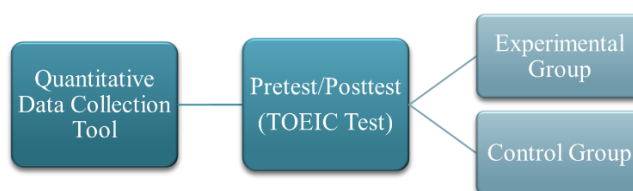
2. A theme-driven, open-ended questionnaire for the experimental group

Two versions of language instruction focusing on speaking skills were devised and executed for the control and experimental groups. The former group was taught via traditional methods and activities, whereas the latter was taught through CRAD model for 12 weeks. The quantitative data were collected using the English proficiency test for speaking skills. Tasks were carefully sequenced, moving from the simple to the challenging levels to elicit optimal linguistic production from learners.

Figure 4.

Phase

One



In the second phase, eighteen information-rich participants were strategically selected from the initial sample of 36 to respond to the open-ended questionnaire. The 18 ESL learners, who were already exposed to the intervention, maintained a 2:1 gender ratio (12 females and 6 males). A 20-item open-ended questionnaire was administered. This instrument was self-crafted based on Affective Filter Hypothesis (Krashen) and Self-Determination Theory (Deci & Ryan), as these theories address the psychological and motivational factors explored in Themes 1 and 3. The Sociocultural Theory (Vygotsky) provided the foundation for Theme 2. The questionnaire was validated through consultation with language specialists and teachers.

Figure 5.

Phase

Two



Pilot Study

The whole process was pilot-tested on 16 ESL students that were not included in the main research. The pilot study provided insights to make changes, such as cutting the length of video clips to 2 minutes and reformulating and rearranging questions in the open-ended questionnaire to add more probing questions. The pilot also verified the functionality of the tools and materials used, pointing out that the open-ended questionnaire provided ease of articulating the thoughts of participants.

Limitations and Delimitations

The study had a number of limitations that affected its scope and generalizability. Primarily, time and the scope of the research, made it necessary to limit the study from other language skills (writing, listening) to speaking only. This practical decision was based on the researchers' unique role in designing and delivering all instructional materials, and focused on the skills that the Pakistani ESL learners most struggle with. A second limitation had to do with the number of participants; it was not possible to find a large enough sample of participants with the desired homogenous characteristics, hence the number of 72 participants out of 160 were selected to form the intervention and control groups (36 participants in each group). Finally, the study did not include the perception of ESL teachers, an insight that was considered important for the further studies that followed the dissemination of the novel CRAD approach. On the other hand, the research also had certain delimitations that determined its scope: it was limited to video as an audiovisual medium for the CRAD intervention. Furthermore, the research site and population were limited to one particular context namely, the English Linguistics Department of The Islamia University of Bahawalpur and specifically undergraduate students of the department as the main reason of accessibility.

Data Analysis

Quantitative data were analyzed using inferential statistics, namely paired and independent sample *t*-tests. The *t*-test was used to check the difference of pre-test and post-test mean score at intragroup comparison as well as intergroup comparison of speaking skills. Qualitative data analysis was done through the six-step thematic analysis proposed by Braun and Clarke (2006) that included familiarization of data, development of initial codes, identification of themes, re-examination of themes, definition and naming of themes, and writing of a final report.

Intragroup Comparison of Pre and Post test Speaking Skill Using Traditional Teaching Method

Table 1.

Mean Scores of Control Group in Pre-test and Post-test (Paired Samples Descriptive Statistics)

Research Group		Mean	N	Std. Deviation	Std. Error Mean
Control Group	Pretest Score in Speaking	67.64	36	25.143	4.191
	Posttest Score in Speaking	107.14	36	17.331	2.888

The results of the paired samples *t*-test of pre-test and post-test scores of the control group are shown in Table 1. The mean of the pre-test was 67.64 (SD = 25.143) while the mean of the post-test was 107.14 (SD = 17.331).

Table 2.
Paired Samples Test

Research Group		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Control Group	Pretest and Posttest Scores in Speaking	-39.500	19.870	3.312	-46.223	-32.777	-11.927	35	.000

P- Value
t- -11.927= P-value .000 is < alpha (.05)

Table 2 presents the difference between the mean scores of the control group in pre- and post-test. The paired-samples *t*-test revealed a mean difference of -39.500, with a 95% confidence interval ranging from -46.223 and -32.777. The *t*-value was -11.927 (*df* = 35) and the *P*-value was 0.000. Since the *P*-value is less than 0.05, the difference in the speaking skills in the control group from the pre-test to the post-test is statistically significant.

Intragroup Comparison of Pre-test and Post-test Speaking Skills Following the Use of Culturally Responsive Audio Description

Table 3.
Mean Scores of Experimental Group in Pre-test and Post-test (Paired Samples Descriptive Statistics)

Group		Mean	N	Std. Deviation	Std. Error Mean
Experimental Group	Pretest Score in Speaking	67.44	36	22.657	3.776
	Posttest Score in Speaking	129.61	36	19.844	3.307

Table 3 presents the speaking skills of the experimental group both before and after the CRAD intervention. The mean pre-test and post-test scores were 67.44 (SD = 22.657) and 129.61 (SD = 19.844), respectively. These descriptive statistics have shown a significant increase in the speaking proficiency of the participants after the CRAD.

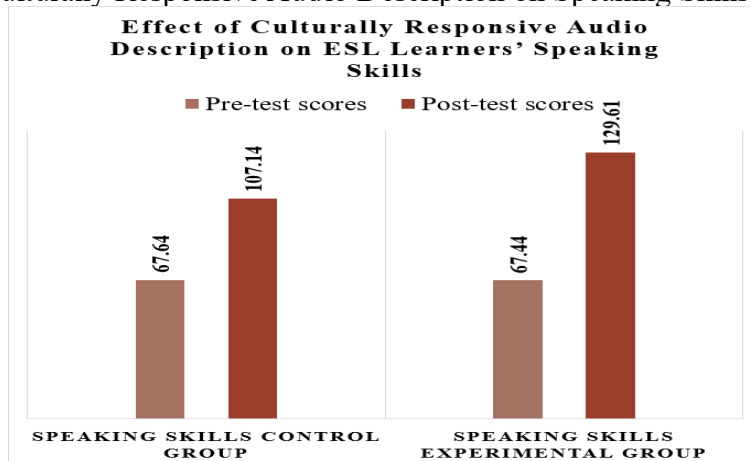
Table 4.
Paired Samples Test

Research Group		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Experimental Group	Pretest and Posttest Scores in Speaking	-62.167	17.369	2.895	-68.044	-56.290	-21.475	35	.000
t=-21.475= P- value .000 is < alpha (.05)									

A paired samples *t*-test was performed to compare the experimental group pre-test and post-test scores. The analysis revealed a mean difference of -62.167, with a 95% confidence interval ranging from -68.044 and -56.290. The *t*-value was -21.475 (*df* = 35) and the *P*-value was 0.000. This result is statistically significant; this means that the CRAD intervention had a great effect in improving the speaking skills of the ESL learners.

Figure 6.

The Effects of Culturally Responsive Audio Description on Speaking Skills



Speaking Skills Intergroup Comparison

Table 5.

Descriptive Statistics (Posttest Mean Scores Control & Experimental Groups)

Groups		N	Mean	Std. Deviation	Std. Error Mean
Posttest Score in Speaking	Control Group	36	107.14	17.331	2.888
	Experimental Group	36	129.61	19.844	3.307

Intergroup comparisons showed that experimental participants, who received CRAD, in the area of instruction of speaking, had significantly higher levels of speaking proficiency in post-test compared to control group participants who received traditional instruction.

Table 6.

Independent Samples Test (Difference in Posttest Mean Scores of Control & Experimental

		Levene's Test for Equality of Variances		t-test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference		Lower	Upper
Posttest Score in Speaking	Equal variances assumed	.768	.384	-5.118	70	.000	-22.472	4.391		-31.230	-13.714
	Equal variances are not assumed			-5.118	68.754	.000	-22.472	4.391		-31.233	-13.712

Group)

An *F*-Test was used to compare the variability in scores of the speaking between the control and experimental group. Subsequently, an independent samples *t*-test was conducted to investigate the effect of the CRAD intervention on speaking skills of the experimental group. The *t*-test value was -5.118 (*df* = 70) and the *P*-value was 0.000, indicating that there was a statistically significant difference between the group means.

Effect Size (Cohen's *d*, 1988)

Table 7.

Effect

Size

Productive Skills	Intervention's Effect size	Description
Speaking Skills	$d=2.919$	A Large Effect Size

The resulting Cohen's *d* value of 2.919 for the CRAD model suggests a very large effect size, indicating substantial practical significance for teaching English speaking skills to ESL learners.

Perceptions of ESL Learners about the Effectiveness of Culturally Responsive Audio Description

The qualitative analysis of learner perceptions was organized around three main themes each of which is based on established pedagogical theory.

Figure 7.
Major Themes



Theme I: Speaking Confidence and Anxiety Reduction (The Local Safety Net)

Questions 1–7 were based on the Affective Filter Hypothesis (Krashen, 1985) directly measuring the reduction of the affective filter that results from incorporating Local Identity-based content.

Theme II: Targeted Linguistic Precision and Complexity

Questions 8-14 were theoretically supported by Sociocultural Theory (Vygotsky, 1978) and Task-Based Language Teaching (TBLT) focusing on the measurement of the cognitive role and efficacy of structured linguistic scaffolding given by the CRAD model.

Theme III: Motivation and Refinement of Speaking Quality

Questions 15-20 were anchored in the Self-Determination Theory (Deci & Ryan, 1985) and had a dual purpose of measuring the resultant effect of the intervention on intrinsic motivation of the learners.

Figure 8.

Theme 1: Speaking Confidence and Anxiety Reduction (The Local Safety Net)

Sub-Themes	Representative Quotes
Content-Driven Anxiety Reduction	"The anxiety dropped because the content was familiar and mine." (Derawar Fort)
Self-Perception as Expert/Guide	"When I describe the Faisal Mosque, I feel like I'm the tour guide." / "I feel more authentic when talking about the Chowk Yadgar."

**Simplicity of Cultural
Transfer**

"Describing how a local dish is prepared felt much simpler... I have the recipe in my mind."

Description of Sub-Theme I: The familiarity of the topic is immediately freeing the cognitive burden of locating content and therefore enables the shift of attention to the language production alone.

Description of Sub-Theme II: Students adopt a professional identity (tour guide, historian) which improves confidence and self-efficacy and supersedes fear of error.

Description of Sub-Theme III: Students find cultural procedures (e.g. cooking or craft making) that they know easier to explain in comparison with foreign abstract topics and therefore easier to transfer from L1 to L2.

Figure 9.

Theme 2: Targeted Linguistic Precision and Complexity

Sub-Themes	Representative Quotes
Precision in Spatial & Relational Language	"Audio Description forces you to be precise with prepositions... beyond the valley or towering over the road." (Karakoram)
Expansion of Lexical Density	"You can't just say 'big'; you must find words like 'vast,' 'expansive,' or 'monumental'..." (Cholistan Desert) / "...words like 'ornate' for the Noor Mahal's interior."
Structural Sophistication	"made me practice using relative clauses ('which is famous for...') to include details quickly and efficiently." (Mohenjo-daro) / "...sequence words like 'first, subsequently, finally' which helps my narrative fluency." (Traditional crafts)

Description of Sub-Theme I: The need to describe the position of objects in a visual way requires the specific use of prepositions and spatial adverbs.

Description of Sub-Theme II: The cultural significance of the content means that students are forced to use more varied and complex adjectives and synonyms, thereby raising the token-to-type ratio.

Description of Sub-Theme III: Sequencing and quick inclusion of details requires the active use of complex sentence structures, i.e. relative clauses and sequencing words.

Figure 10.

Theme 3: Motivation and Refinement of Speaking Quality

Sub-Themes	Representative Quotes
Pride as a Refinement Driver	"I practice harder when the topic is Pakistani culture because it is a topic of national pride. I want my speaking to be refined to represent my identity well."

Authoritative Delivery

"I feel I am speaking with more authority... This makes me use a stronger voice and better pronunciation..." / "I noticed my intonation and pacing improved..."

Purposeful Language Use	"The AD task makes my English feel important... This purpose drives me to perfect my descriptive skills."
--------------------------------	---

Description of Sub-Theme I: The link to national or cultural pride creates an inner desire to achieve a higher level of proficiency to represent accurately their identity.

Description of Sub-Theme II: Confidence in the cultural content equates to better paralinguistic contents, such as voice control, pacing and intonation.

Description of Sub-Theme III: Students see the task as a communication act of meaning (where a global language is used to explain local identity) rather than a simple classroom task.

Discussion

To increase the efficacy of this mixed-method research, the current study combined data findings of both phases. The quantitative results have been supported by the qualitative ones not only helped in confirming the internal validity of the study, but also offered strong evidences in support of the alternative hypothesis.

Research Question 1: Does the use of the Culturally Responsive Teaching-Audio Description (CRAD) model lead to a statistically significant improvement in the speaking skills (fluency, linguistic complexity, and coherence) of undergraduate ESL learners compared to a conventional teaching method?

The quantitative findings provide a conclusive answer. The experimental phase revealed that the CRAD group's overall speaking proficiency was significantly higher ($p < .05$) across all measured metrics—fluency, linguistic complexity (T-unit analysis), and coherence—compared to the control group. This finding supports the theoretical assumption that the CRAD model is an effective linguistic scaffold. The particular needs of the Audio Description (AD) process, the need to convey the complicated visual data as a brief and real-time speech, forced learners to actively use the high-level language that they require to progress. The density of descriptive vocabulary and complex syntax had the most significant improvements in the CRAD group.

The literature on the AD is a robust corroboration of this finding since it acknowledges the ability of the method to propel linguistic accuracy and enhance descriptive lexis in the ESL setting (Orero & Wharton, 2007; Perego & Trevisan, 2024). In addition, such a systematic direction towards the enhancement of particular aspects of the language (Theme 2: Targeted Linguistic Precision and Complexity) is consistent with the current level of agreement regarding the effectiveness of contextual, interactive means to improve speaking strategies in the Pakistani setting (Bango et al., 2023; Yasmeen et al., 2023). The fact that the observed improvement confirms that the requirement of the CRAD model that is structured and task-based effectively converted the affective security offered by local content into objective and measurable improvements in advanced speaking competency (Javaid et al., 2024a,b).

Research Question 2: What are the undergraduate ESL learners' reflections regarding how the integration of local cultural content influences their speaking confidence, anxiety levels, and motivation to refine their spoken language ("Global Voice")?

The qualitative data provides a significant understanding of the process through which the substantial quantitative enhancement was achieved, offering strong evidence supporting the influence of local identity as an affective and motivational factor. The results of the reflections of the undergraduate ESL learners provide a comprehensive description of the degree to which the CRAD model meets the requirements of these learners in terms of confidence, managing anxiety and motivation.

A. Speaking Confidence and Anxiety Reduction (The Local Safety Net)

Students in overwhelming numbers echoed that when local Pakistani heritage material was used (such as descriptions of places like the Noor Mahal or Derawar Fort) a Local Safety Net effect was achieved, which reduced speaking anxiety immensely. This observation is a direct testimony to Affective Filter Hypothesis (Krashen, 1985). Although students said their thoughts were clear to find a good adjective, as stated in Theme 1 (Speaking Confidence and Anxiety Reduction (The Local Safety Net)), they claimed that their mind had been made free of thoughts since the content had already been mastered. The intervention helped to provide a low-anxiety environment required to be engaged in risk-taking and effective practice as it lowers the cognitive load associated with unfamiliar topics. This is exactly the same outcome as in the Pakistani context by Amjad et al. (2024) who found that there is an increase in confidence and fluency when familiar content is used and by Idrus et al. (2023) who found that CRT leads to less shy and more active speakers.

B. Theme 3: Motivation and Refinement of Speaking Quality

The qualitative data showed that the connection of the task with cultural identity was a strong intrinsic motivational factor. As stated in the Theme 3, namely, Motivation and Refinement of Speaking Quality, the learners shared the opinion that they saw themselves as more authoritative speakers because they took the role as the experts of the culture. This feeling of national and cultural pride created a need to have a perfect pronunciation and formal vocabulary to make sure that their message was delivered in a perfect and accurate tone using a Global Voice. Consequently, this strong association between the identity and communicative purpose made the task not a simple academic task but an empowering social action. This intrinsic motivation mechanism is highly supported by Jalil (2023) and Reskyani et al. (2024) whose findings in the context of Malaysian and Indonesian higher education revealed that the CRT had a significant effect on motivation and willingness to improve speaking skills. The results of this study do not exhibit any contradictory findings since they validate already existing literature that culturally relevant pedagogy has a positive influence on student engagement and motivational outcomes.

Conclusion & Implications

The present study definitively demonstrates the effectiveness of the CRAD model as an

effective pedagogical approach for developing the speech competence of undergraduate ESL learners. The results effectively confirm one of the key theoretical frameworks that are organized into three pillars, namely the Affective Gateway (Krashen), the Motivational Engine (Deci & Ryan), and the Linguistic Scaffold (Vygotsky & TBLT). The quantitative data, which indicated statistically significant improvements ($p < .05$) in fluency, linguistic complexity and coherence, were highly supported by qualitative data that indicated that exploiting Local Identity lessens anxiety (The Affective Gateway) and generates the intrinsic drive required to produce more complex linguistic products (The Linguistic Scaffold) and to attain an authoritative global voice. The research concludes that the CRAD model is a strong, evidence-based pedagogic model that has sweeping implications on practice. It highlights a fact that the best language learning takes place when the matter of subject legitimizes the identity of the learner and turns the task to be an assessment of lack to an expression of empowered speech. Hence, learning institutions should incorporate culturally responsive and task-based interventions, such as CRAD in undergraduate education to improve advanced oral skills and confidence that result in academic and professional achievement.

CRAD Implementation Recommendations

For curriculum designers and institutions, the main recommendation is to legitimize and validate the CRAD model. This begins with the requirement of the incorporation of Culturally Responsive-Audio Description Tasks into undergraduate-level courses in ESL based on local heritage and content. This confirms the Local Identity aspect of the student and establishes the aspect of confidence that is fundamental in academic advancement. To make this part institutionalized, the CRAD model as a fundamental pedagogical instrument to form certain high-order speaking skills, including a linguistic complexity (e.g., advanced T-units and rich descriptive lexis) and coherence, should be adopted by the institutions in question. Such adoption necessitates strategic resource creation, i.e. the investment in a collection of high-quality, short video clips depicting Pakistani landmarks and cultural practices for use in licensed in-class Audio Description assignments.

In the case of ESL teachers, the emphasis is placed on the tactical application to ensure the maximum benefit of the model (affective and linguistic). The teachers need to deliberately refocus the content to form through the familiar cultural content to reduce the affective filter by shifting the cognitive burden of knowing *what to say* to *how exactly can I describe this*. At the same time, the educators must maximize the role of the expert, and therefore design the activities in a way that continuously places the ESL learners within the role of a cultural expert or a tour guide. This improves the self-efficacy and authoritative voice, and paralinguistic aspects required in the Global Voice. Lastly, educators need to directly instruct in descriptive lexis and syntax, spending the time in the classroom directly teaching the aspects of the language needed in Audio Description, including the use of complex prepositional phrases, relative clauses, and advanced adverbs, to get students beyond the level of fluency and towards the level of precision in their speech.

Future Research Directions

To extensively prove and optimize the CRAD pedagogical model, three primary areas should be the subject of research in the future. First, longitudinal studies are necessary to examine the long-term transferability of the linguistic improvements (complexity and coherence) gained with CRAD, assessing their viability when applied to other challenging communicative tasks such as academic presentations, argumentation, and impromptu debate. Second, the CRAD model's efficacy must be replicated across a wide range of cultural environments to determine the extent to which its success is dependent on the specific cultural context under consideration versus its universal pedagogical framework, thereby defining its generalizability. Third, studies are required to determine the efficacy of teacher training programs directly focused on the

application of the CRAD model, specifically examining the effect of such training on teacher self-efficacy, adherence to the model, and ultimately, student achievement.

References

- Amjad, S., Sarwat, S., Asif, A., & Farooq, M. A. (2024). ESL learners' perceptions about the effect of language preferences on English learning outcomes. *Journal of Asian Development Studies*, 13(4), 941-956.
- Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. K. (2009). *Introduction to research in education* (8th ed.). Wadsworth.
- Bango, Z. A., Channa, S., & Niaz, P. (2023). Factors influencing speaking and listening skills of ESL learners at university level in Sindh. *Pakistan Journal of Humanities and Social Sciences*, 11(4), 4251-4260.
- Bausells-Espín, A. (2022). Audio description as a pedagogical tool in the foreign language classroom: An analysis of student perceptions of difficulty, usefulness and learning progress. *Journal of Audiovisual Translation*, 5(2), 152-175.
- Bennett, S., Gunn, A., Gayle-Evans, G., Barrera, E., & Leung, C. (2018). Culturally responsive literacy practices in an early childhood community. *Early Childhood Education Journal*, 46, 241-248. <https://doi.org/10.1007/s10643-017-0839-9>.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>.
- Burger, G. (2016). Audiodeskriptionen anfertigen – Ein neues Verfahren für die Arbeit mit Filmen [Creating audio descriptions - A new method for working with films]. *Info DaF-Informationen Deutsch als Fremdsprache*, 43(1), 44-54.
- Bygate, M. (1987). *Speaking*. Oxford University Press.
- Caingcoy, M. (2023). Culturally responsive pedagogy: A Systematic Overview. *Diversitas Journal*, 8(4), 3203-3212. <https://doi.org/10.48017/dj.v8i4.2780>
- Calduch, C., & Talaván, N. (2018). Traducción audiovisual y aprendizaje del español como L2: El uso de la audiodescripción [Audiovisual translation and learning Spanish as an L2: The use of audiodescription]. *Journal of Spanish Language Teaching*, 4(2), 168-180. <https://doi.org/10.1080/23247797.2017.1407173>
- Cenni, I., & Izzo, G. (2017). Audiodescrizione nella classe di italiano L2: un esperimento didattico [Audiodescription in the Italian L2 classroom: A teaching experiment]. *Incontri*, 31(2), 9-24.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed-methods approaches* (2nd ed.). SAGE.
- Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, 19(2), 109-134.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of teacher education*, 53(2), 106-116.
- Green, J. C., & Caracelli, V. J. (Eds.). (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*. Jossey-Bass.
- Gunn, A. A., Bennett, S. V., Alley, K. M., Barrera IV, E. S., Cantrell, S. C., Moore, L., & Welsh, J. L. (2021). Revisiting culturally responsive teaching practices for early childhood preservice teachers. *Journal of Early Childhood Teacher Education*, 42(3), 265-280.
- Herrero, C., Sánchez-Requena, A., & Escobar, M. (2017). Una propuesta triple: Análisis fílmico, traducción audiovisual y enseñanza de lenguas extranjeras [A three-

- pronged approach: Film analysis, audiovisual translation and foreign language teaching]. In *inTRAlinea Special Issue: Building Bridges between Film Studies and Translation Studies*. <http://www.intralinea.org/specials/article/2245>
- Ibáñez-Moreno, A., & Vermeulen, A. (2013). Audio description as a tool to improve lexical and phraseological competence in foreign language learning. In D. Tsigari & G. Floros (Eds.), *Translation in language teaching and assessment* (pp. 41–61). Cambridge Scholars Publishing.
- Ibáñez-Moreno, A., & Vermeulen, A. (2014). La audiodescripción como recurso didáctico en el aula de ELE para promover el Desarrollo integrado de competencias [Audiodescription as a didactic resource in the ELE classroom to promote the integrated development of competences]. In R. Orozco (Ed.), *New directions on Hispanic linguistics* (pp. 263–292). Cambridge Scholars Publishing.
- Ibáñez-Moreno, A., & Vermeulen, A. (2015a). Using VISP (VIdeos for SPeaking), a mobile app based on audio description, to promote English language learning among Spanish students: A case study. *Procedia - Social and Behavioral Sciences*, 178(2), 132–138.
- Ibáñez-Moreno, A., & Vermeulen, A. (2015b). Profiling a MALL app for English oral practice: A case study. *Journal of Universal Computer Science*, 21(10), 1339–1361.
- Ibáñez-Moreno, A., & Vermeulen, A. (2015c). VISP 2.0: Methodological considerations for the design and implementation of an audio-description based app to improve
- Idrus, F., Ramli, L. N., & Habib, N. J. (2023). Exploring Preservice Teachers' Experiences of Implementing Culturally Responsive Teaching in the ESL Classrooms. *Theory & Practice in Language Studies (TPLS)*, 13(3), 766–776.
- Jalil, C. R. A. (2023). The application of culturally and linguistically responsive pedagogy in english speaking classrooms—A case study. *International Journal of English Language Teaching*, 11(2), 72–91.
- Javaid, Z. K., Ramzan, M., & Ijaz, S. (2024a). A systematic review on cognitive and motivational impact on English language learning through artificial intelligence. *International Journal of Literature, Linguistics and Translation Studies*, 4(1), 44–71.
- Javaid, Z. K., Chen, Z., & Ramzan, M. (2024b). Assessing stress causing factors and language related challenges among first year students in higher institutions in Pakistan. *Acta Psychologica*, 248, 104356.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
- Kroll, T., Neri, B. A. M. (2009). Designs for mixed-methods research. In S. Andrew & E. J. Halcomb (Eds.), *Mixed-methods research for nursing and the health sciences* (pp. 31–49). Blackwell Publishing Ltd.
- Ladson-Billings, G. (2021). Three decades of culturally relevant, responsive, & sustaining pedagogy: What lies ahead? *The Educational Forum*, 85(4), 351–354.
- Leaver, B. L., & Willis, J. R. (Eds.). (2004). *Task-based instruction in foreign language education: Practices and programs*. Georgetown University Press.
- Littlewood, W. (1981). *Communicative language teaching*. Cambridge University Press.
- Maszerowska, A., Matamala, A., & Orero, P. (Eds.). (2014). *Audio description: New perspectives illustrated*. John Benjamins Publishing Company.
- Matiso, N. H. (2024). Optimising culturally responsive pedagogies in multicultural English second language classrooms. *International Journal of Learning, Teaching and Educational Research*, 23(11), 384–401.
- Mehta, A. (2024). Culturally responsive teaching: Bridging the achievement gap. *Global International Research Thoughts*, 12 (1), 49–54.
- Menacho-Vargas, I., Garcia, U. C., Rojas, M. L. R., Wong-Fajardo, E. M., & A Saavedra-

- López, M. (2021). The effect of culturally-based conversations on developing speaking skill among Peruvian upper-intermediate EFL learners. *International Journal of Society, Culture & Language*, 9(3), 109-118.
- Morrison, S. A., Brown Thompson, C., & Glazier, J. (2022). Culturally responsive teacher education: Do we practice what we preach?. *Teachers and Teaching*, 28(1), 26-50.
- Navarrete, M. (2013). El doblaje como herramienta de aprendizaje en el aula de español y desde el entorno Clipflair [Dubbing as a learning tool in the Spanish classroom and from the Clipflair environment]. *MarcoELE*, 16, 75–87.
- Nawaz, S., Aqeel, M., Ramzan, M., Rehman, W., & Tanoli, Z. A. (2021A). Language, Representation and Ideological Stance of Brahui In Comparison with Urdu and English Newspapers Headlines. *Harf-O-Sukhan*, 5(4), 267-293.
- Nawaz, S., Aqeel, M., & Ramzan, M. (2021B). Listening Comprehension Problems, Corresponding Factors and Strategies for Better or Enhanced Listening Skill. *Pakistan Languages and Humanities Review*, 5(2), 729-737.
- Nawaz, S., Sharif, S., Ramzan, M., & Nawaz, H. (2022). The portrayal of women in billboard advertisements installed in Bahawalpur, Punjab, Pakistan. *Harf-o-Sukhan*, 6(1), 200-207.
- Nisar, M., Parveen, S., & Asif, M. (2023). Analyzing the role of socio-cultural norms in Pakistani English language education: A case study of ESL classroom practices. *Panacea Journal of Linguistics & Literature*, 2(2), 89-103.
- Nosheen, S. S., Javed, M., & Akhtar, H. (2020). Analyzing strategies for developing students speaking skills in public universities of Pakistan. *Global Regional Review*, 5(1), 92-99.
- Nunan, D. (2003). *The learner-centered curriculum*. Cambridge University Press.
- Orero, P., & Wharton, S. (2007). The audio description of a Spanish phenomenon: Torrente 3. *The Journal of Specialised Translation*, (7), 164–178. http://www.jostrans.org/issue07/art_orero_wharton.pdf
- Palmer, E. (2014). *Teaching the core skills of listening and speaking*. ASCD.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). SAGE.
- Perego, E. (Ed.). (2014). *L'Audiodescrizione Filmica per I Ciechi e gli Ipovedenti* [Filmic Audiodescription for the Blind and Visually Impaired]. EUT.
- Rafi, M. S., & Fox, R. K. (2021). Translanguaging as a culturally responsive pedagogy for teaching English to multilingual Pakistani university students. *SPELT Quarterly Journal*, 36(1), 1–18.
- Reskyani, R., Dollah, S., Basri, M., & Mustaqimah, M. (2024). Students' motivation in learning speaking using culturally responsive teaching approach in higher education. *EXPOSURE Yapedumen: Publikasi Jurnal Ilmiah Akademik Universitas Muhammadiyah Makassar*, 13(2), 387-398.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge University Press.
- Rodrigues Barbosa, E. (2013). La audiodescripción como herramienta para la enseñanza de ELE [Audio description as a tool for teaching Spanish as a foreign language]. In L. M. A. de Freitas (Ed.), *Anais do xiv congresso brasileiro de professores de espanhol: Língua e ensino* (Vol. 1, pp. 487–496). PEERJ.
- Rubin, A., & Babbie, E. R. (2010). *Essential research methods for social work*. Brooks/Cole.
- Samiullah, M. (2018). Effect of teaching through songs on speaking skills in English among early graders of Pakistan. *Journal of Early Childhood Care and Education*, 2, 51–60.
- Savignon, S. J. (1991). Communicative language teaching: State of the art. *TESOL Quarterly*,

25(2), 261–278.

- Smith, T., Avraamidou, L., & Adams, J. D. (2022). Culturally relevant/responsive and sustaining pedagogies in science education: Theoretical perspectives and curriculum implications. *Cultural Studies of Science Education*, 17(3), 637-660.
- Talaván, N., & Lertola, J. (2016). Active audiodescription to promote speaking skills in online environments. *Sintagma*, 28, 59–74.
- Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.
- Warren, C. A. (2018). Empathy, teacher dispositions, and preparation for culturally responsive pedagogy. *Journal of Teacher Education*, 69(2), 169-183.
- Wesley-Nero, S., & Donley, K. (2024). Culturally and linguistically responsive pedagogy: Examining teachers' conceptualizations of affirmative instructional practices for multilingual learners. *TESOL Journal*, 15(4), e881.
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.
- Yasmeen, S., Fareed, M., & Naeem, A. (2023). Teaching English language speaking skills to business students at tertiary level: Practices in Pakistan. *Bahria University Journal of Humanities & Social Sciences*, 6(1), 153-168.
- Zengin, B., & Yıldız, T. A. (2024). Metaphor analysis of audio description in English language teaching.

The Literacy Trek, 10(1), 103–130.

Appendix: Screenshots of Culturally Responsive-Audio Description Tasks



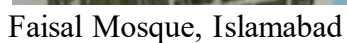
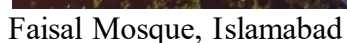
Derawar Fort, Bahawalpur

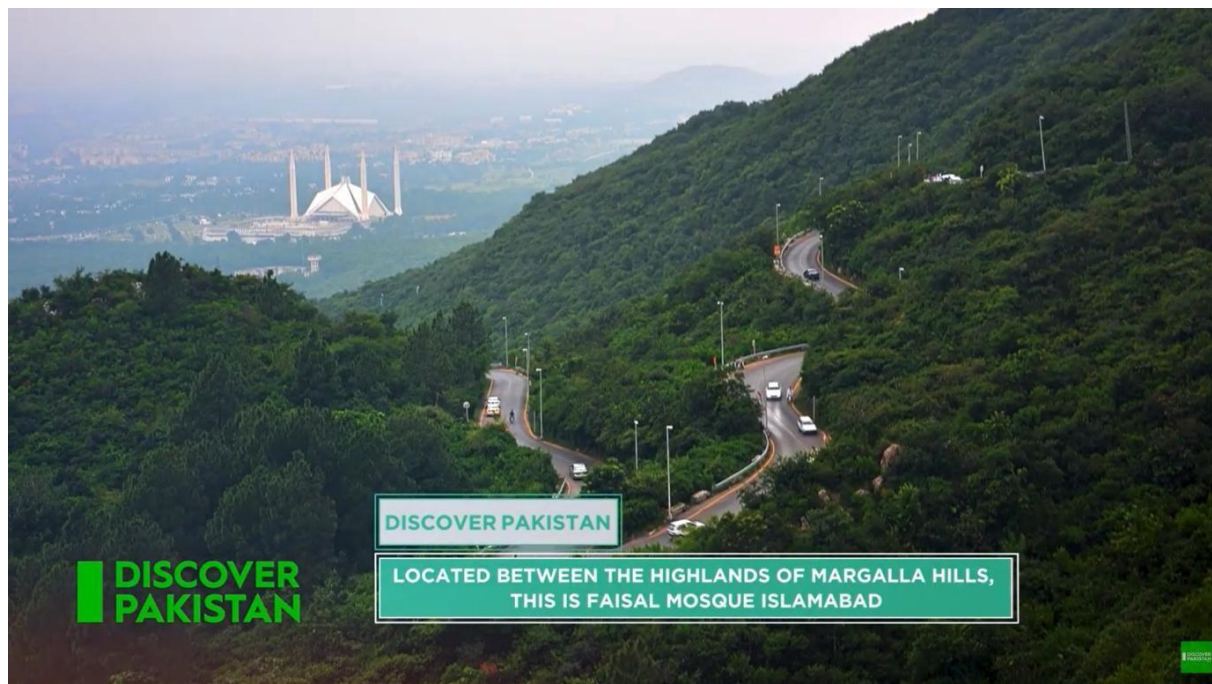


Noor Mahal, Bahawalpur



Shalamar Garden, Lahore





Margalla Hills, Islamabad



Bab-e-khyber, KPK



Mazar-e-Quaid (The National Mausoleum), karachi

Culture of Pakistan

As students were enrolled from diverse remote areas, the content deliberately included cultural elements from all regions of Pakistan. Moreover, the commitment to Local Identity made in this study was thorough and inclusive and acknowledged that the heterogeneity of the student body included members of the Hindu community whose cultural heritage was also embedded into the culturally responsive materials.



Celebrating Eid Event



Peshawari Qehwa: A Taste of Tradition



Punjab's Famous Food: Saag, Makki Roti, Lassi



The Traditional Sindhi Female Dress Code



The Traditional Sindhi Ajrak



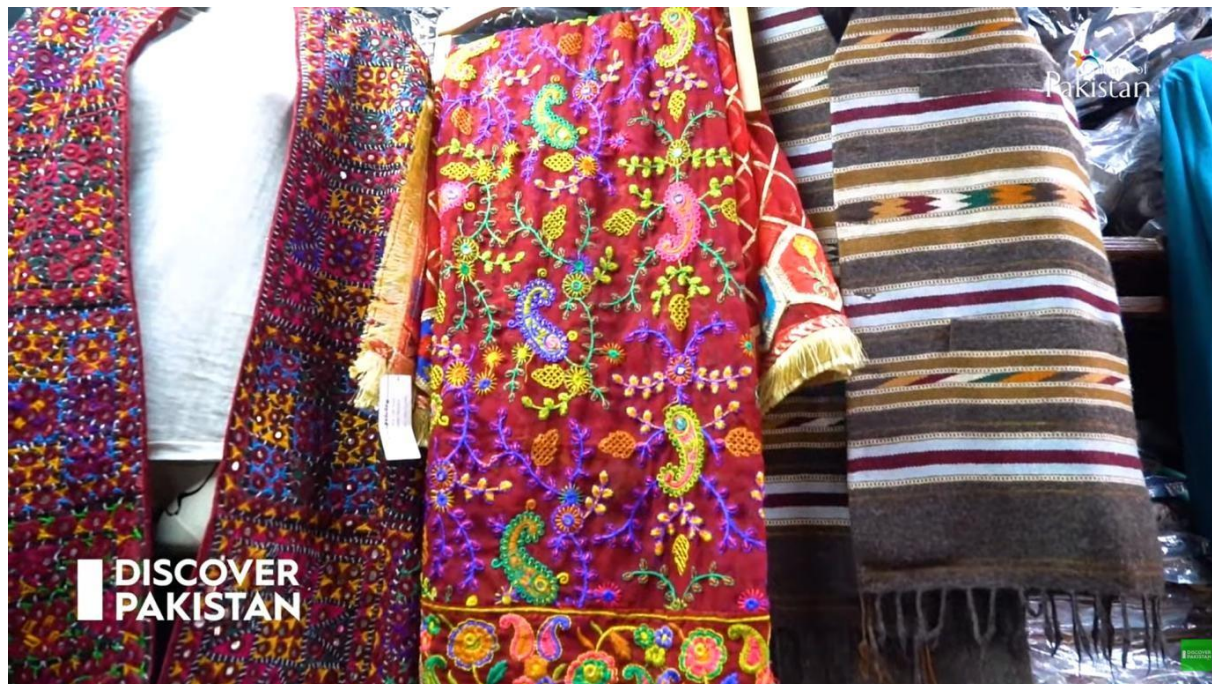
Kpk Rhythm: Rubab (Rabab)



Polo in Balochistan



Balochi Male Attire



Kashmiri Female Attire



Kashmiri Chai (Pink Tea)



The Beautiful Baltistani Female Caps



Neza Bazi (Tent Pegging) in Punjab



Kabaddi in Punjab



Sajji: The Balochi Feast



Namkeen Rosh: The Balochi Feast



Balochi Female Dress Code



Balochi Male Traditional Attire



Gilgit/Baltistan Dress Code



The design of Traditional Jewelry from Azad Jammu & Ka

In pictures: Diwali in Pakistan



Women of Hindu community are lighting lamps on *rangoli* patterns. Photo: OINN

In pictures: Diwali in Pakistan



Members of Hindu community are lighting sparklers in Karachi. Photo: OINN
shmir