People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

University of Kasdi Merbah Ouargla

Faculty of Letters and Languages

Department of Letters and English Language

Thesis submitted to the Department of Letters and English Language in Candidacy for the Degree of "Doctorate Es-science" in English Language and Literature

Major: Applied Linguistics

Title

The Effects of Embedded Multimedia on Learners' Pronunciation

The Case of First-Year Students of English at Biskra University

Submitted by

Walid AOUNALI

Supervised by

Prof. Saliha CHELLI

Board of examiners

Mohammed Seghir HALIMI	Prof	University of Ouargla	Chair
Saliha CHELLI	Prof	University of Biskra	Supervisor
Nawal DIB	Ass Prof (MCA)	University of Ouargla	Examiner
Djalal Eddine AMRANI	Ass Prof (MCA)	University of Ouargla	Examiner
Wafa AGGOUN	Ass Prof (MCA)	University of Batna 2	Examiner
Abir Soundous GHASKIL	Ass Prof (MCA)	University of Batna 2	Examiner

2025-2026

Title

The Effects of Embedded Multimedia on Learners' Pronunciation

The Case of First-Year Students of English at Biskra University

Submitted by

Walid AOUNALI

Statement of Authorship

I, **Walid AOUNALI**, do hereby declare that the present doctoral thesis contents titled

The Effects of Embedded Multimedia on the EFL Learners' Pronunciation: The Case of First-Year Students of English in Biskra University is my proper and original research work supervised by **Professor Saliha CHELLI**. I do respect and follow the chart of ethics of research integrity and the academic policy of the University of Ouargla. All the sources used or quoted in this thesis are properly indicated in references section.

Researcher: **Walid AOUNALI**

Date: 04/12/2025

Acknowledgments

First and foremost, I would like to express my sincere gratitude and appreciation to my supervisor, **Prof. Saliha CHELLI**, who introduced me to the world of speech sounds and the English pronunciation realm, for her invaluable guidance, encouragement, and unwavering support throughout the journey of my research study. Her extensive expertise and insightful feedback have been influential in shaping both this thesis and my professional development as a researcher and a mentor in teaching English pronunciation.

I am also grateful to the esteemed members of the jury for agreeing to examine my thesis; particularly, **Prof. Mohammed Seghir HALIMI** for accepting to preside the panel of examiners as well as the respected jury members for dedicating their efforts and time in reviewing and scrutinising the thesis; namely, **Dr. Nawal DIB**, **Dr. Djalal Eddine AMRANI**, **Dr. Wafa AGGOUN** and **Dr. Abir Soundous GHASKIL** for their constructive feedback and insightful comments to enhance the final version of my thesis.

My heartfelt thanks go to my colleagues in the Departments of English language and Literature in both universities of Ouargla and Biskra, whose collaboration made this journey less daunting. Specifically, **Dr. Abdelhak CHENINI** and **Mr. Hicham FELLAH** as well as **Mr. Mahmoud REZIG**.

Special thanks go to **Dr. Yasser BENMOUSSA**, **Dr. Khaled LEBIAR** and **Prof. Abderezzek BEN ZAOUI** for their assistance with coding and statistical analysis of the collected data. I am truly grateful to all those who have contributed in a way or another in realizing this thesis mainly first-year students who participated in the study.

The researcher: Walid AOUNALI

Dedication

This thesis is dedicated:

To my family, for their faithful prayers and encouragement that pushed me to carry on this long journey yet a worthwhile experience in the academia.

To my former teachers, for shaping who I am today. Additionally, my sincere gratitude is due to all those who believed in me, even when I fainted, this achievement is accomplished thanks you.

ABSTRACT

This doctorate thesis investigates the effects of embedding digital materials under the framework of multimedia-based instruction in teaching English pronunciation and their impact on the Algerian EFL learners. Specifically, it adopts mixed-method research approach that analyses the speech productions of 20 first-year students of English at Biskra University, alongside questionnaire data collected from 201 respondents. In this regard, the study examines whether multimedia resources enhance the potentials of EFL learners to better pronounce English language or not. Accordingly, the participants were divided into two groups following a quasi-experimental design over eight weeks, totalling twelve hours of intensive instruction targeting segmental-level and sentence-level proficiency. The scores were compiled for statistical comparison with the native-speaker audio recordings using the acoustic analysis software *Praat* in terms of formants' means, range and standard deviation to measure statistical differences for hypothesis testing and validation. The results revealed significant statistical differences in favour of the alternative hypothesis (H1) at (Spearman's $\rho = .306–.415$; $p < .001$; Cronbach's $\alpha = .812$) showing improvements in the participants' pronunciation gains post-intervention. Key findings demonstrate that embedding multimedia tools in pronunciation instruction not only enhance learners' pronunciation and intelligibility but also leads to measurable progress in their ability to produce clear and comprehensive English speech. The study recommends the incorporation of multimedia-based instruction in the EFL syllabi and highlights some problematic areas for investigation in future research on the impact of technology-based language learning and teaching in the age of AI.

Keywords: Multimedia-based instruction; pronunciation; EFL learners; segmental and suprasegmental features.

RESUME

Cette recherche doctorale traite du rôle de l'intégration des contenus numériques multimodaux dans l'enseignement de la prononciation en classe d'anglais langue étrangère (ALE) à l'université Mohamed Khider à Biskra. Pour ce faire, deux outils d'investigation ont été adoptés : l'analyse des productions orales et l'administration d'un questionnaire. Dans cette approche mixte, nous avons recueilli les productions orales de 20 étudiants du premier cycle, et nous avons administré un questionnaire à 201 apprenants. Dans le cadre d'un dispositif expérimental, les participants ont été répartis en deux groupes (groupe expérimental et groupe témoin). Pendant huit semaines, les participants ont suivi un enseignement intensif portant sur la maîtrise segmentale et la maîtrise au niveau de la phrase. Les scores ont été ensuite collectés et comparés à des enregistrements audios de locuteurs natifs à l'aide du logiciel d'analyse acoustique PRAAT afin de mesurer les différences statistiques pour le test et la validation des hypothèses. Les résultats révèlent des différences significatives en faveur de l'hypothèse alternative (H1) avec un score de (Spearman ρ = .,306–,415; $p < ,001$; Cronbach α = ,812). Les conclusions de cette étude montrent que l'intégration d'outils multimédias dans l'enseignement de la prononciation améliore non seulement la prononciation et l'intelligibilité des apprenants, mais entraîne également des progrès mesurables dans leur capacité à produire un discours clair et compréhensible en anglais. Ces résultats pourraient offrir des implications pédagogiques efficientes pour l'enseignement de l'anglais en Algérie et dans des contextes similaires, suggérant que l'utilisation stratégique des ressources multimédias peut remédier efficacement aux difficultés de prononciation et contribuer au développement global de la compétence communicative.

Mots-clés : Multimédia; prononciation; apprenants d'anglais langue étrangère (ALE); traits segmentaux et suprasegmentaux.

الملخص

تتناول رسالة الدكتوراه توظيف المحتوى الرّقمي من خلال التّعليم المعتمد على الوسائط المتعدّدة في تدريس نطق اللغة الإنجليزية، لإيجاد مدى تأثير ذلك على متعلّمي الإنجليزية كلغة أجنبية في الجزائر. كمقاربة بحثيّة تعتمد منهج تحليل الإنتاج الشفوي لعشرين طالبًا جامعيًا في تخصّص اللغة الإنجليزية بجامعة محمد خيذر بسكرة، بالإضافة إلى استبيان شمل 201 مشاركًا. في هذا الإطار، يتقصّى البحث ما إذا كانت موارد الوسائط المتعدّدة تعزّز قدرات متعلّمي الإنجليزية على تحسين نطق اللغة الإنجليزية. وعليه، قُسّم المشاركون إلى مجموعتين في تصميم شبه تجريبي دام ثمانية أسابيع، باثنتي عشرة ساعة من التّدريس المكثّف على المستوى المقطعي للأصوات و فوق المقطعي للجمل. بعد ذلك، جُمعت النتائج لمقارنتها بالتسجيلات الصوتية للمتحدّثين الأصليين، وذلك باستخدام برنامج «برات» للتحليل الصوتي من حيث معدّلات صيغ الصوائت ومداها، والانحراف المعياري من أجل قياس الفروق اللإحصائية في اختبار الفرضيّات والتحقّق منها. وبالمثل، دُعّمت هذه التحليلات بإجراءات إحصائية لبيانات استبيان الطلبة حول التّعليم المعتمد على الوسائط المتعدّدة. كشفت النتائج عن وجود فروق ذات دلالة إحصائية لصالح الفرضية البديلة ($\alpha = 0,812$) و معامل كرونباخ و بما يُظهر تحسّنًا في نطق المتعلمين بعد تطبيق الدراسة التجريبية. Spearman's $\rho = .306–.415$ وتبرهن النتائج على أنّ إدماج أدوات الوسائط المتعدّدة في تدريس النطق لا يعزّز نطق المتعلّمين وفهم كلامهم فحسب، بل يؤدّي أيضًا إلى تقدّم ملموس في قدرتهم على تحدث واضح ومفهوم بالإنجليزية و لهذه النتائج دلالات تربوية مهمّة لتعليم اللغة الإنجليزية في الجزائر وسياقات مماثلة، إذ تشير إلى أنّ الاستخدام الاستراتيجي لموارد الوسائط المتعدّدة يمكن أن يُسهم بفاعلية في معالجة صعوبات النطق، ويعزّز الكفاءة التّواصلية الشاملة. وتوصي الدّراسة بضرورة إدماج التّعليم القائم على الوسائط المتعدّدة في مناهج الإنجليزية، مع إبراز بعض المجالات المقترحة للبحث المستقبلي في تعلّم اللغات المدعوم بالتكنولوجيا في عصر الذكاء الاصطناعي.

الكلمات المفتاحية: التعليم بالوسائط المتعدّدة ،متعلمي الإنجليزية ، النطق ، الخصائص القطعية و الفوقطعية.

List of Abbreviations and Acronyms

Abbreviations	Full Form
AI	Artificial Intelligence
ASR	Automatic Speech Recognition
CALL	Computer-Assisted Language Learning
CAPT	Computer-Assisted Pronunciation Teaching
CBA	Competency-Based Approach
CG	Control Group
EG	Experimental Group
ELT	English Language Teaching
L1	First Language
L2	Second Language
MBI	Multimedia-Based Instruction
NS	Native Speaker
NNS	Non-Native Speaker
Praat	Speech Analysis Software
RP	Received Pronunciation
SLA	Second Language Acquisition
SPSS	Statistical Package for the Social Sciences
TPACK	Technological Pedagogical Content Knowledge
VOT	Voice Onset Time
ZPD	Zone of Proximal Development

List of Tables

List of Figures

Table of Contents

Chapter one:

General Introduction

General Introduction

Introduction

The present doctoral study seeks to investigate EFL learners' perceptions and attitudes toward the integration of multimodal approach to develop their fluency and communicative competences to reach native-speakerism which is regarded an artificial goal that the vast majority of learners cannot achieve in non-natural settings of exposure to the target language (Saito, 2021) in real-life situations.

The teaching of English pronunciation in the EFL contexts creates a challenge in pronunciation pedagogy. Despite recent pedagogical innovations in teaching and learning, the EFL learners find it difficult to accomplish intelligible and comprehensible pronunciation, particularly in the Algerian context where exposure to authentic input is very limited, and large class sizes, in addition to outdated syllabi dedicating traditional methods of teaching and overlooking pronunciation. In Algeria, these difficulties are related to phonological interference and lack of awareness, and reliance on teacher-centred as well as drill-based methods without effective practice or immediate corrective feedback.

1.1. Background of the Study

In the 21st century, English has increased use and propagation as the lingua franca of international communication, science, technology, and commerce (Jenkins, 2000; Crystal, 2012). This dominance has added unprecedented demand for effective English language teaching (ELT) worldwide, particularly in non-native contexts where learners must master English to communicate internationally and keep up with scientific innovations. However, the complex phonological systems of English for non-native speakers (NNS) created a language barrier for many EFL learners to achieve communicative competence. In Algeria, where English was

introduced in the educational curricula after the independence as a second foreign language together with to Arabic as the national language of teaching. In addition, French has a strong position in education due to colonial heritage as a language of administration and teaching in some disciplines like medical studies and scientific branches. This is why, the teaching of pronunciation has historically been marginalised, often restricted to elementary drills within phonetics modules (Benrabah, 2014; Bouhadiba, 2018) and the total absence of audiovisual aids in the textbooks of the Ministry of National Education except for some phonics sections dedicated to transcription and pronouncing final “s” and “ed” that most teachers avoid in teaching (Aounali, 2014). Still, as the new linguistic policy in the country that started prioritising English as a Medium of Instruction in tertiary level and other sectors to align with global economic demands. (as in Ministry of Higher Education and Scientific Research directives in 2022).

In digital age, multimedia technologies have revolutionised ELT by bridging gaps in authentic input and interactive practice (Chapelle, 2003; Mayer, 2014). Tools such as high-fidelity videos, speech-recognition applications, and acoustic visualisation software enable learners to access unlimited native and non-native models, receive immediate feedback, and engage in self-paced refinement features that are absent in traditional Algerian classrooms characterised by large groups (40–50 students) and limited resources. As grounded in Cognitive Theory of Multimedia Learning (Mayer, 2009), these technologies exploit dual-channel processing (visual & auditory) to reduce cognitive load and enhance retention, particularly for adult learners challenging with fixed errors from L1/L2 interference (Flege, 1995).

Despite these previous studies, empirical research on multimedia-based instruction instruction (MBI) on EFL learners' pronunciation in the Arab and North African contexts remains scarce, with most studies focusing on perceptions of students rather than statistical acoustic evidence (Ghounane, 2018; Bengrait, 2020). In Algeria, where learners face unique trilingual challenges and interference (Arabic-French-English), there is a lack of empirical studies that systematically look into embedding MBI into university curricula to show its effects on segmental accuracy and suprasegmental fluency. This thesis addresses that void, examining how MBI can transform pronunciation from a neglected skill into a cornerstone of communicative competence in the Algerian EFL context.

1.2. Statement of the Problem

From our own experience, we have noticed that the students are scarcely trained on phonetic drills and listening and repeating activities or using authentic audio or video material for improving their pronunciation. English pronunciation is not taught as separate skill or subject matter but we can find it within the module of phonetics or oral expression, which makes students' ability to pronounce correct English utterances futile. The results of this study will undoubtedly reveal the effects of embedding multimedia tools in improving pronunciation.

Despite growing interest in using authentic materials in teaching pronunciation in the digital age with the availability of native-speakers' model and even non-native models to illustrate the correct pronunciation on the internet or offline to improve learners' pronunciation from the era of recorded audio tracks and educational videos to the age of the internet and the rise of Automatic Speech Recognition (ASR) tools like Siri and others for providing instant feedback and corrections like Duolingo and BetterAccent Tutor to the recently breakthrough innovations in the AI-Driven tools like Siri and others for pronunciation training.

The knowledge gap exists in our context but prevailing research works conducted among Arab in general and the Algerian EFL learners in specific remain limited in some key areas. For that, various studies have either lacked a robust and practical strategies on how to implement the materials, tools and applications for improving students' outcomes, that is, they failed to confirm that all the participants successfully benefitted to draw realistic generalisations about the value of such tools. Additionally, the conceptual frame in tracking the research gap seeks to compensate the observed weaknesses in previous studies to explore learners' deficiency in pronunciation.

The present study has systematically scrutinised learners' perceptions of use of multimedia-based instruction in improving their pronunciation to address the existing gaps in the Algerian context by employing this instructional method so as to get more reliable data and offer practical implementations. Besides, by exploring this gap, we are interested in finding an answer to both deficiency in EFL learners' pronunciation outcomes and their attitudes and preferences in improving their overall fluency.

Algerian EFL university students consistently consider pronunciation as a source of communicative difficulty and anxiety, yet current teaching practices fail to address this need effectively. Moreover, traditional methods rely exclusively on teacher imitation, occasional minimal-pair drills, and written transcription exercises or the exposure to connected speech which is not enough without corrective feedback. So, a variety of pronunciation errors affect the intelligibility of learners' confidence in oral production. On the international scale, recent research supports the efficacy of technology-enhanced pronunciation training (Levis, 2007; Derwing & Munro, 2015; Saito, 2021). However, empirical studies in the Algerian tertiary context remain scarce and the existing work is mainly descriptive or limited to measuring attitudes. Moreover, the

explicit contribution of systematic use of multimedia materials is scarcely conducted- has not been rigorously examined.

2. Significance of the Study

This doctoral project explored how to improve English pronunciation teaching at university level. More specifically, we look into the extent to which English pronunciation is improved through digital content and how it can foster foreign language learners' level in accuracy and fluency in order to improve their pronunciation through constant training using multimedia materials. Since, pronunciation is one of the most resilient challenges in second language acquisition, particularly for adult EFL learners in contexts where exposure to authentic English is limited and classroom time is dominated by grammar rules and written tasks.

In Algeria, first-year and second-year university students regularly exhibit persistent segmental errors and suprasegmental mistakes (monotonous tone, misplaced word stress, long pauses and slow rhythm) that significantly reduce intelligibility and comprehensibility for native listeners. These difficulties are rooted in lack of exposure to English in instruction, strong phonological interference from Arabic and French, overcrowded classrooms, and a near-exclusive reliance on teacher-centred methods of teaching, handout-based lessons that provide neither sufficient models of authentic speech nor systematic corrective feedback.

Despite the rapid global proliferation of multimedia and mobile technologies capable of delivering unlimited native-speaker input, real-time acoustic visualisation, and automatic speech recognition, Algerian university curricula have yet to integrate such tools in any structured, syllabus. Pronunciation continues to be treated as a marginal sub-component of "listening and speaking" or "phonetics" modules with little or no use of video modelling, Praat-based self-

analysis, or AI-driven applications that have been shown internationally to accelerate perceptible gains even in short interventions (Derwing & Munro, 2015; Levis, 2020; Saito, 2021). This disconnect constitutes a critical research and pedagogical gap: while learners and teachers increasingly recognise pronunciation as a priority, no empirical study has yet examined the systematic implementation and measurable impact of multimedia-based pronunciation instruction within the specific constraints and affordances of the Algerian tertiary context.

3. Aim of the Study

The primary aim of this doctoral research is to determine whether the intervention of syllabus-embedded multimedia pronunciation programme can produce statistically significant improvements in segmental accuracy, suprasegmental features, and overall intelligibility among first-year EFL students at Mohamed Khider University of Biskra, as well as surveying EFL learners' perceptions on the effectiveness of multimedia in teaching pronunciation. The study ended up with several recommendations for improving learners' pronunciation and to apply more effective methods in order to enhance the practice in English pronunciation teaching in our universities. Teaching pronunciation mainly concentrates on phonetics instruction in which individual sounds and separate words stress placement instead of focusing on suprasegmental features of speech and fluency. Using multimedia resources and authentic language materials in communicative tasks that explicitly focus on pronunciation will be more effective in improving students' pronunciation. In fact, using multimedia resources in the classroom would allow teachers to model native-speakers pronunciation clearly to their students in an effective way.

3. Research Questions

The present study addresses the pedagogical and empirical gaps identified in the literature review. The formulated research questions are based on recent studies, and the problems that the Algerian EFL students face, which emphasise the need to test both measurable phonetic improvements and the EFL learners' perceptions. So, we seek to answer the following questions:

1. Does the use of multimedia tools in pronunciation instruction lead to significant improvements in learners' fluency and accuracy compared to traditional methods?
2. Is embedding multimedia resources in teaching and learning English language pronunciation effective and motivating to learners?
3. How does a multimedia-based syllabus for teaching English pronunciation facilitate both teachers' instruction and students' engagement with authentic materials?
4. What are the potential challenges and difficulties in learners' pronunciation and how to overcome these discrepancies in intelligibility and comprehensibility of their foreign accent?

1.4. Hypotheses

As far as the English pronunciation is concerned, the present study tests the use multimedia resources to teach pronunciation from different aspects. Primarily, the use of authentic tools in order to make students aware of the segmental features of English and suprasegmental features can be exemplified as well. The study tests the following paired hypotheses:

Null Hypothesis (H_0): The use of multimedia-based instructional resources has no statistically significant effect on students' pronunciation performance.

Alternative Hypothesis (H_1): The use of multimedia-based instructional resources has a statistically significant positive effect on students' pronunciation performance.

1.6. Methodology

This research study was carried out in the university of Biskra in Algeria to investigate the effectiveness of multimedia-based learning and teaching materials on improving students' pronunciation. In other terms, whether EFL students prefer to use multimedia materials over the traditional paradigm to improve their pronunciation. To achieve the purpose of our research study, we have opted for mixed-methods design as a research methodology. More than that, the actual investigation used different instruments of data collection. Hence, this study involved questionnaires to survey students' viewpoints about teaching and learning pronunciation in general and the use of multimedia in particular. Besides, an experimental study was conducted to have a cross-sectional view about the current study.

Through this study, a quasi-experimental design was opted by the researchers in order to analyse the effects of multimedia-based instruction as the independent variable with reference to pronunciation improvement the dependent variable. The main reason behind this experiment is to find an answer to the question: Could multimedia resources employed in the classroom enhance the potentialities to better pronounce English language? The students who participated in our research study to was divided into 2 groups. The first group has to learn English pronunciation using multimedia-based instruction, whereas the other group has to learn English pronunciation using the same content and teaching method, but without embedding multimedia in the classroom. The subjects have had some test and recorded while pronouncing English on *word-level proficiency test (phoneme-level)* and *sentence-level proficiency test* (prosodic-level) before the experimentation (pre-test) and after treatment during the different stages of the experiment (post-

test). Both tests utilised the same wordlists and sentences in which we started with word-level then sentence-level afterwards then analysed acoustically using the speech analysis software Praat.

1.7. Data Analysis

The participants of the current study were chosen in a simple random selection from the whole population of First-year students at Biskra university in which we have selected 20 students to be subjected to the experimental study, in turn, they were divided into a control group during a period of eight weeks during which data were gathered and recorded. In addition, as a supporting instrument, we administered a questionnaire in a purposeful representative sampling of 201 first-year students to survey their viewpoints about the multimedia tools we have used and about their expectations and experiences. The collected data were analysed quantitatively using SPSS version 26 software and Praat speech analysis software. In practice, the data of this study will be analysed acoustically for finding differences among the score of the pre-test and once in the post-test. The recordings measured vowel formants (F1, F2), voice onset time (VOT), word stress duration, and intonation contour. Then, the audio files were checked again by another teacher to make sure the measurement was correct. Afterwards, the recordings were subjected to human rater scores from 1 to 5 for “how easy the speech is to understand” and “how strong the accent is” as advised in the work of (Derwing & Munro, 2015). In addition, the questionnaire responses were analysed using to calculate means and percentages and check reliability and to find correlations (Spearman’s test).

8. Structure of the Thesis

The present thesis explores the integration of multimedia tools in teaching English pronunciation to Algerian EFL learners, emphasising educational theories, experimental design, and practical implications for both learners and teachers. The current study is organised into five

chapters; in the first chapter, we set out the general introduction as a roadmap to the study outlining the role of multimedia in language education in previous studies. Then, the statement of the problem identifying key challenges, including the research gap in the Algerian context. The second chapter then, focusing on the review of the main literature in the field of ELT and second language pronunciation synthesising theoretical and contextual insights into language learning and pronunciation instruction. In the third chapter, the methodology section details the framework of the research design adopting mixed-methods approach and data analysis as well as ethical considerations. For the fourth chapter, it is dedicated to the discussion of the results of both the empirical study and the questionnaire in order to find any statistical differences in students' outcomes and to explore learners' perceptions of pronunciation instruction, then, we discuss the research questions and hypotheses and we conclude it with summary of the findings. Finally, the last chapter, concludes the study with the pedagogical implications and contributions addressing teacher readiness in multimedia integration and the learners' issues in pronunciation improvement. In the end, we proposed a model for multimedia-enhanced pronunciation pedagogy to end the study with some prospects on the future of multimedia use in learning and teaching including the use of AI for pronunciation training.

Chapter Two:

Review of

Literature

Chapter Two: Review of Literature

Introduction

The present chapter sets out the basic concepts of language learning theories from the beginnings of this century to the recent contemporary educational theories, technological integration, and evolving pedagogies. Starting from foundational learning theories like behaviourism, cognitivism, and constructivism then arriving at the scope of our study on pronunciation skills. More specifically, the integration of multimedia in language pedagogy is reinforced by dual coding theory and the cognitive theory of multimedia learning (CTML), which postulates that learning is enhanced when information is presented through both verbal and visual channels.

In the current chapter, we are going to review the pronunciation instruction and its recent paradigm that has shifted from a focus on native-like accuracy to prioritizing intelligibility and comprehensibility. As shown, meta-analyses reveal that instruction targeting both segmental and suprasegmental features demonstrates more improvements in comprehensibility than in nativelikeness. Subsequently, there is a necessity of using technology-enhanced tools such as speech recognition software and corpus-based resources which provide real-time feedback and authentic materials. In specific EFL contexts like Algeria, historical, sociolinguistic, and multilingual factors complicate English pronunciation instruction.

In this regard, teachers and learners face several defies include interference from Arabic and French phonological systems, as well as the lack of teachers' training, infrastructure and authentic educational resources. In the conclusion, the review of literature showcased a contrastive analysis of the three languages currently used in Algeria to address these issues in multilingual learners.

2.1. Revisiting Learning Theories in the 21st Century

Prior to embarking on the recent learning theories in the EFL context, we will have a synopsis of the traditional theories that shaped the foundations of modern learning theories. Thus, it is necessary to revisit the main classical theories of learning such as behaviourism, cognitivism, humanism, and constructivism in light of today's modern theories that shifted from teacher-centred approach to learner-centred, lifelong learning and problem-solving approaches that reflect the 21st century skills, and rapidly changing world. Finally, we conclude the literature review with the basic concepts of multimedia use in the teaching and learning of foreign languages with a special reference to the place of multimedia-based instruction in the EFL milieu.

The rapid evolution of technology, globalisation, and the availability of knowledge all over the world have considerably reshaped the landscape of language learning in the 21st century. Before examining modern approaches, it is essential to revisit the classical theories that continue to influence modern pedagogy today. In the past, behaviourism, cognitivism, humanism, and constructivism, were major theories in teaching and learning, whereas they are now considered as basic principles that clarify different components of the learning process (Schunk, 2020; Siemens, 2005).

In the digital age, Siemens (2005) and Downes (2012) proposed connectivism as a new theory for networked learning. For them, knowledge no longer resides just in the individual but in connections between nodes (people, devices, databases). In pronunciation pedagogy, this manifests as learners creating personal content on networks for learning: YouTube channels, pronunciation forums, speech-recognition apps, and peer voice-note exchanges, all of which extend learning far beyond classroom settings. These classical and modern theories have mostly progressed language education from teacher-centred approach to learner-centredness, lifelong

learning, technology-mediated construction of knowledge (Richards & Rodgers, 2014; Kern, 2015). All in all, the 21st-century classroom is then featured by:

- Blended learning and flipped classroom models,
- Task-based language learning and teaching and project-based learning,
- Focus on autonomous learning and digital learning,
- Integration of authentic multimodal materials.

All of these changes are particularly applicable to pronunciation, a component of speaking that needs considerable exposure to target language, corrective feedback, and repeated practice, knowing that traditional teaching methods cannot provide these features (Derwing & Munro, 2015;

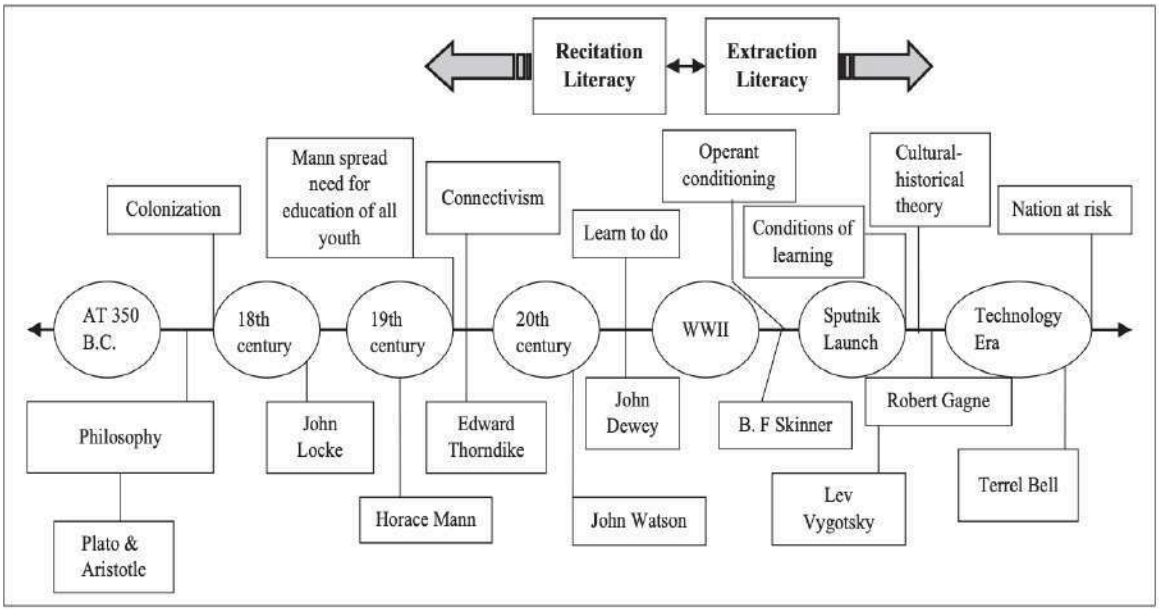
2.1.1. Behaviourist Approaches to Language Learning

Behaviourism emerged in the early 20th century as one of the first systematic attempts to explain learning through observable behaviour. This approach, pioneered by practitioners such as John B. Watson, Edward Thorndike, and namely B.F. Skinner, views learning as a process of habit formation resulting from stimulus-response. Within this framework, internal mental processes are largely disregarded, as behaviourists focus on measurable changes in learners' actions rather than on their cognitive states (Skinner, 1953). In language learning, the behaviourist perspective became highly influential during the 1950s and 1960s, particularly through the Audio-Lingual Method (ALM). This method was grounded in behaviourist psychology and structural linguistics, emphasizing repetition, mimicry, and reinforcement as means to develop language proficiency.

Learners were expected to acquire correct pronunciation and grammatical structures through continuous drills and positive reinforcement, leading to the formation of accurate linguistic habits (Richards & Rodgers, 2014). Errors were viewed as undesirable and were immediately corrected to prevent the establishment of faulty language patterns.

Figure 1.1

Literacy timeline from Early Philosophy to Present



Adapted from (Edgar, 2012, p. 5)

From the behaviourist perspective, pronunciation is a matter of *conditioning*. Learners imitate native models and receive immediate feedback or correction from the teacher, in the hope of reinforcing accurate pronunciation. For example, repeated exposure to a specific phoneme such as the non-Arabic sounds in English /ŋ/ followed by positive feedback strengthens the learners' ability to produce it correctly. This conditioning process aligns well with the principle of **operant conditioning**, where reinforcement (e.g., praise, good marks, or digital badges) increases the possibility of accurate pronunciation behaviours recurring (Skinner, 1957).

Table 1

Skinner's operant therapy techniques

Techniques	Effect on Behaviour	Examples
Positive reinforcement	Increase	£5 given by teacher when done the homework, merit marks, praise, sweets for participation in class.
Negative reinforcement	Increase	£5 need to give to the teacher if failed to finish homework given, Withdrawal of threat of punishment if a pupil behaves well in class.
Punishment - produce something unpleasant	Decrease	Give students a detention when they are not behaving correctly in class.
Punishment - remove something pleasant	Decrease	Prevent child from participate in class activity such as play games.

Note. adapted from Burhanuddin, Ahmad, Said, & Asimiran, 2021, p. 88

In pronunciation teaching, behaviourism translated into endless choral and individual repetition of sounds, words, and sentences after a native-speaker model (often from tape recorders). Immediate correction and over-learning were believed to "stamp in" correct articulation (Celce-Murcia et al., 2010).

Although heavily criticised from the 1970s onward for ignoring meaning, cognition, and creativity (Chomsky, 1968; Richards & Rodgers, 2014), behaviourist principles remain alive and highly effective in specific areas such as:

- Automaticity of basic pronunciation patterns,

- Mastery of difficult individual sounds through focused drilling.

Despite its contribution to early language teaching methodologies, behaviourism has been criticized for its mechanical view of learning and its neglect of mental and cognitive processes. However, its principles remain intact in modern EFL classrooms, especially in technology-assisted pronunciation training. Multimedia tools that provide instant feedback, such as speech recognition mobile applications and pronunciation computer software programs, still employ behaviourist principles by reinforcing correct responses and discouraging incorrect ones. Hence, while traditional drill-based methods may seem outdated, their digital adaptations continue to influence pronunciation pedagogy in the multimedia age of today.

2.1.2. Cognitivist Approaches to Language Learning

Cognitivism appeared in the 1960s and 1970s as a direct reaction to behaviourism. Instead of seeing learners as passive receivers of stimuli, cognitivists view them as active information processors who perceive, organise, store, and retrieve knowledge (Atkinson & Shiffrin, 1968; Piaget, 1970).

Cognitivism emerged ultimately as a reaction to the limitations of behaviourism, shifting the focus from observable behaviour to internal mental processes namely perception, memory, and problem-solving. Influenced by scholars like Jean Piaget, Jerome Bruner, and David Ausubel, the cognitivist approach views learning as an active process of information processing, where learners construct knowledge through understanding rather than habit formation (Schunk, 2020).

Constructivism as paradigm or worldview hypothesize that learning is active and constructive process. Information received by the learner is constructed throughout the journey of learning. According to this theory, people will actively construct knowledge according to their own understanding about reality. New information is linked to a previous

> knowledge thus mental representation are subjective. (Burhanuddin, Ahmad, Said, & Asimiran, 2021, p. 88)

In language learning, cognitivist theories emphasize that learners develop mental representations of linguistic structures and rules, which they use to produce and comprehend language. Learning a language, therefore, involves more than imitation—it requires internalizing phonological, lexical, and syntactic systems. Piaget's theory of cognitive development suggests that learners progress through stages of cognitive growth, while Bruner's discovery learning emphasizes the role of active engagement and scaffolding in language acquisition. Both ideas underscore the importance of meaningful interaction and comprehension in language learning (Bruner, 1960).

In pronunciation teaching, cognitivism brought three major changes:

- Explicit explanation of how sounds (tongue position, voicing, aspiration) instead of blind repetition.
- Use of visual support (mouth diagrams, waveforms, spectrograms) to help learners create accurate mental models.
- Focus on meaningful practice rather than mechanical drills.

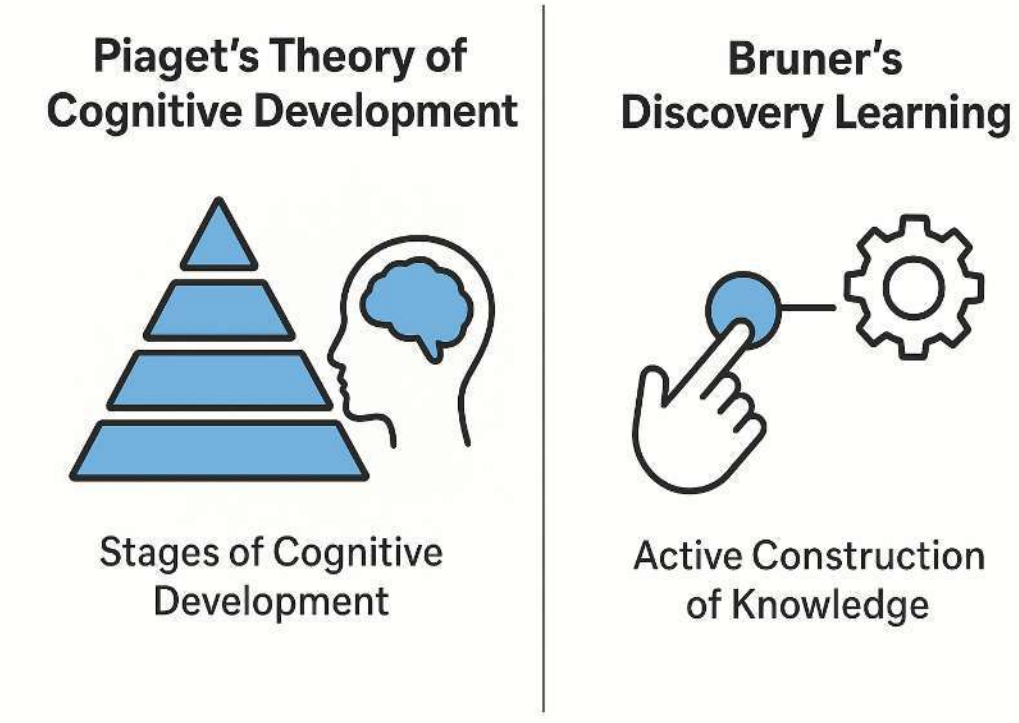
This explains why a short video showing lip rounding and productions of sounds is much more effective than only hearing the sound, which uses both channels without overloading them. Although cognitivism is sometimes criticised for being too "in the head" and ignoring social factors, it works perfectly together with multimedia tools and remains the main theory behind most successful pronunciation apps used by millions of learners worldwide (Levis, 2018).

The most important cognitivist contribution to multimedia pronunciation teaching is Richard Mayer's Cognitive Theory of Multimedia Learning (CTML) (2009, 2014). Mayer's three principles are now standard:

- Dual-channel principle, that is, humans process visual and verbal information through separate channels.
- Limited-capacity principle means each channel can only handle a small amount at a time.
- Active-processing principle is the real learning that happens when learners select, organise, and integrate new information with what they already know.

Figure 1.2

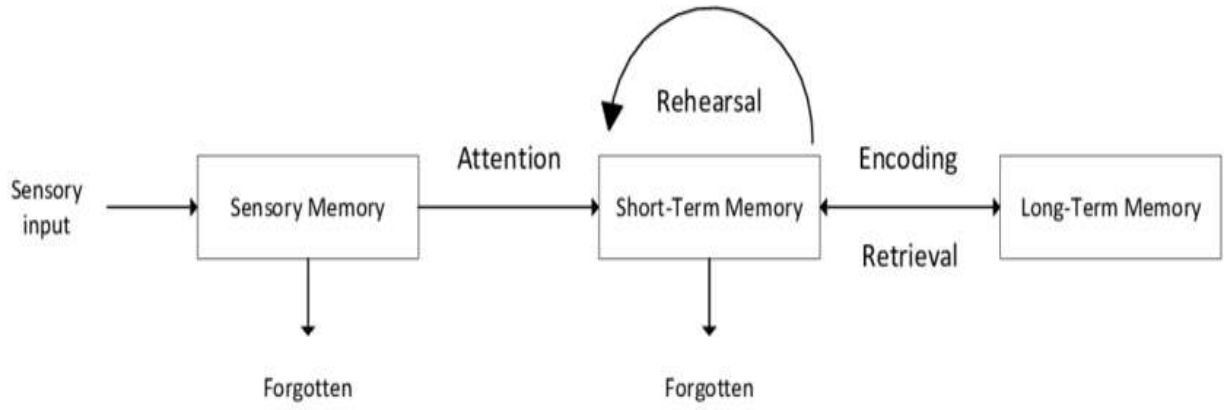
Cognitive Development to Discovery Learning: Piaget and Bruner



When applied to pronunciation learning, cognitivism highlights the learner's awareness and understanding of phonological rules. Instead of mere repetition, learners analyse how sounds are produced, considering articulation features like tongue position or airflow, and consciously apply this knowledge during practice. Information processing models (Atkinson & Shiffrin, 1968) explain how auditory input is received, processed, stored, and retrieved, providing a framework for understanding how multimedia tools can enhance pronunciation learning. For example, animated diagrams of mouth movements and synchronized audio-visual input help learners encode and retain phonetic information more effectively.

Figure 1.3

Information processing model



Note. Adapted from Samrgandi, 2023, p. 112

Richard Mayer's Cognitive Theory of Multimedia Learning (2009) integrates these principles by suggesting that learners process information through dual channels (visual and auditory) with limited capacity. When well-designed, multimedia materials reduce cognitive overload and promote deeper learning. Thus, in EFL pronunciation teaching, interactive videos, spectrograms, and pronunciation visualizers support cognitive engagement by linking perception

with production. Cognitivism provides a more comprehensive understanding of language learning as a mental and strategic process. It bridges theory and practice by explaining how learners internalize language patterns and how technology can facilitate this internalization. Unlike behaviourism's external reinforcement, cognitivism values learners' internal motivation, awareness, and reflection: these are the essential qualities for mastering pronunciation in multimedia EFL classroom.

2.1.3. Schema Theory and Information Processing

Within the cognitivist framework, Schema Theory and Information Processing Theory provide crucial insights into how learners perceive, organize, and retain linguistic information. Schema theory, introduced by Bartlett (1932) and later developed by Rumelhart (1980), posits that knowledge is stored in mental frameworks or "schemas" that guide comprehension and production. In language learning, schemas help learners interpret new input by connecting it to existing linguistic and conceptual knowledge. For instance, when learning new pronunciation patterns, learners rely on prior phonological schemas to categorize unfamiliar sounds. This process of assimilation and accommodation (terms borrowed from Piaget) illustrates how learners continuously reshape their internal representations as they encounter new linguistic data.

In practice, Schema Theory and Information Processing are two key parts of the cognitivist approach that help us understand how learners handle new pronunciation information. When we hear a new sound or word, the brain tries to fit it into an existing schema (top-down processing) or creates a new one (bottom-up). Once the schema is built, the sound becomes automatic and easy to produce and recognise. So, the Information Processing Model (Atkinson & Shiffrin, 1968) proposed that the brain works in three stages:

1. Sensory memory (hears the sound for a second).

2. Short-term / working memory (can hold only 5–9 items for 20–30 seconds).

3. Long-term memory (stores forever if practised enough).

Pronunciation problems happen when too much new information goes to working memory at once then it will be overload and nothing reaches long-term memory. So, multimedia can help (Sweller's Cognitive Load Theory, 1988; Mayer, 2009) through:

- Giving information in small pieces (segmenting principle)
- Using pictures and sound together (modality principle)
- Removing unnecessary words or decoration (coherence principle)

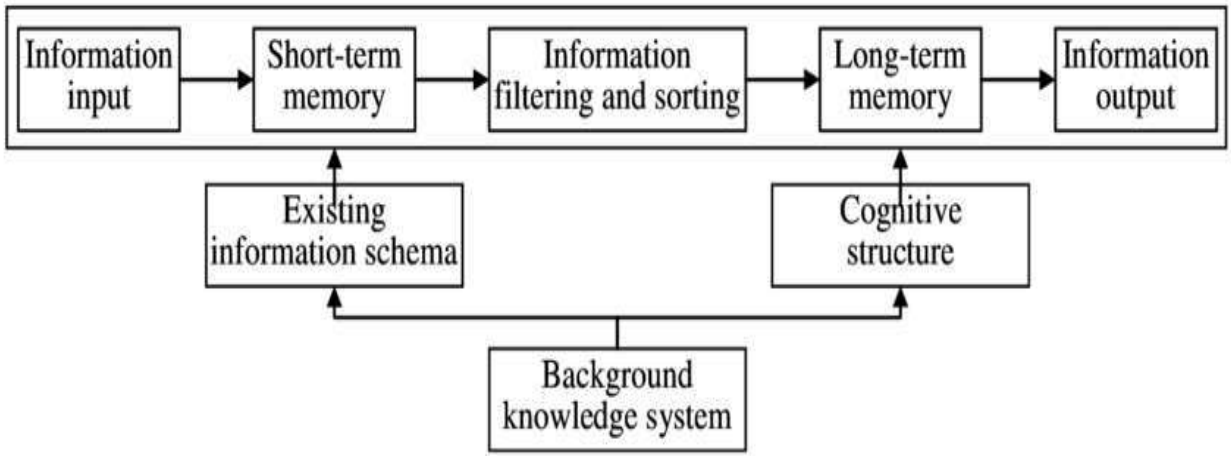
Essentially, to reduce working-memory load and help the correct schema move quickly to long-term memory learners can follow these examples:

- Instead of only saying "Listen and repeat: ship – sheep", we can show:
- a short video of mouth shape
- waveform in Praat
- written word + picture of a ship/sheep
- instant recording and comparison

In short, schema theory explains why learners mix up sounds (wrong mental file), and information processing explains how multimedia helps fix it (less load, better storage). Together they are the main reason why short, visual, multimodal practice works so well for adult pronunciation learners (Mayer, 2014; Levis, 2018). Thus, learners can avoid the same mistakes again and again through good information processing with multimedia and the teacher tells them how to fix it quickly using authentic models from natives.

Figure 1.4

Cognitive process of schema theory



Note. Adapted from (Ma, 2016, p. 567)

Information Processing Theory complements schema theory by describing how information moves through distinct stages: sensory input, short-term memory, and long-term memory (Atkinson & Shiffrin, 1968). For EFL learners, pronunciation practice involves encoding audio input, processing it through short-term memory, and transferring it into long-term memory through repetition and meaningful practice. Multimedia tools can enhance this process by providing multiple inputs: visual (lip movement diagrams), auditory (native pronunciation), and textual (phonetic symbols), which supports deeper encoding and recall. Thus, schema activation and efficient information processing are central to multimedia pronunciation learning.

2.1.4. Cognitive Theory of Multimedia Learning (CTML)

Building on these ideas, Richard Mayer's Cognitive Theory of Multimedia Learning (CTML) (2009) provides a comprehensive explanation of how learners integrate information from words and pictures. CTML is based on three key principles:

(1) people possess dual channels for processing verbal and visual input,

(2) each channel has limited capacity,

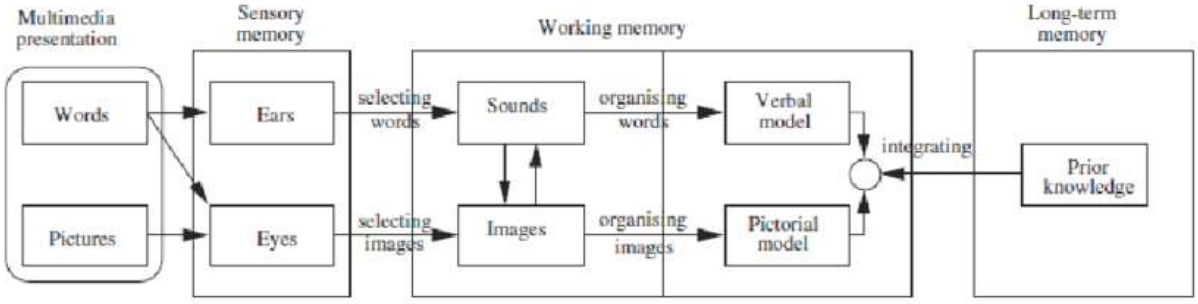
(3) meaningful learning occurs through active processing-selecting, organizing, and integrating new information with prior knowledge.

In the context of pronunciation learning, CTML explains how combining auditory and visual cues enhances comprehension and retention. For example, when learners watch a video demonstrating mouth movements while simultaneously hearing the correct pronunciation, they process the information through both channels, leading to more effective learning. However, CTML also warns against cognitive overload. Therefore, effective multimedia design should present information clearly, limit distractions, and guide learners' attention to essential elements (Mayer, 2014).

In practice, Mayer's research shows that lessons designed with these rules improve learning by (60–120%) compared to normal teaching and this means that learners reach clear, understandable speech much faster. Subsequently, the use of short clips without distractions in addition to spoken explanation with pictures, and immediate self-comparison. For him, real learning happens only when the learner selects, organises, and connects new information to what they already know.

Figure 2

Mayer's Cognitive Theory of Multimedia Learning



Note. Adapted from Mayer, 2009

Together, Schema Theory, Information Processing, and CTML form the cognitive support of multimedia pronunciation learning. They demonstrate that learning occurs most effectively when information is organized, connected to prior knowledge, and presented in a balanced multimodal format that supports both comprehension and memory retention. In the same regard, Mayer tried to find an answer to the question "How does the human brain learn when we give it words and pictures together?" and after more than 30 years and 200 experiments, he provided three practical rules in learning.

1. Dual-Channel Principle: We have two separate channels: one for words (ears and reading) and one for pictures (eyes).

2. Limited-Capacity Principle: Each channel can only hold 4–7 pieces of new information at the same time.

3. Active-Processing Principle: Learning happens only when the learner **selects, organises, and connects** the new information next to prior knowledge.

2.1.5. Constructivist Approaches to Language Learning

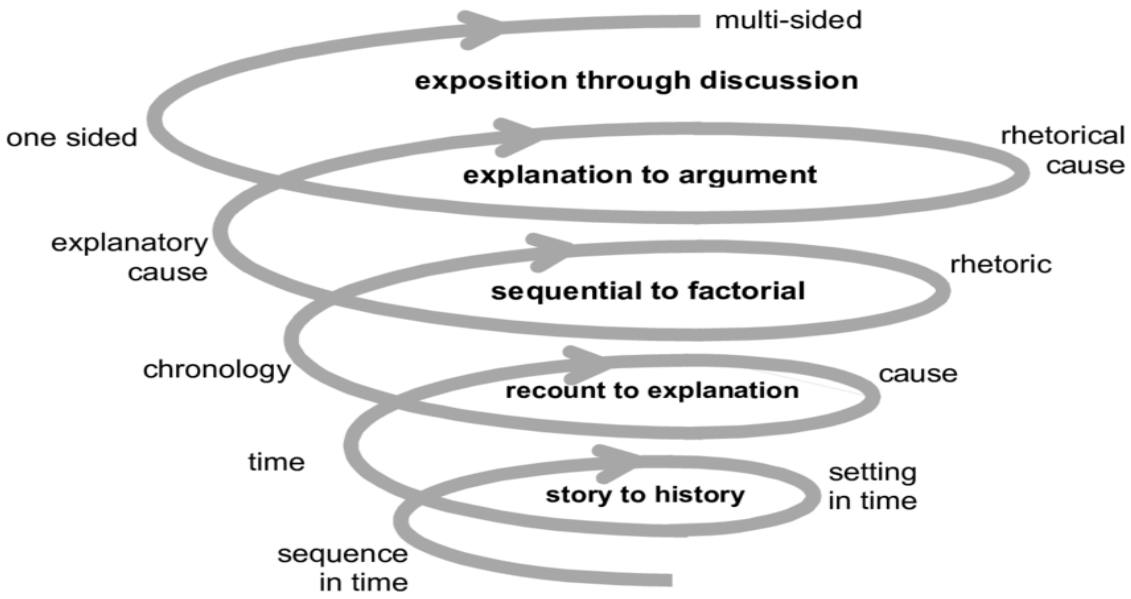
Constructivism emerged as a response to the limitations of behaviourism and cognitivism, emphasizing the learner's active role in constructing meaning rather than passively receiving information. As Schunk (2020) explains, "Constructivism emerged as a reaction to behaviourist theories that viewed learning as a passive process of stimulus–response conditioning. In contrast, constructivists argue that learners actively construct their own understanding and knowledge of the world, through experiencing things and reflecting on those experiences" (p. 231). This perspective shifts the focus from teaching as transmission to learning as active exploration. Jean Piaget and Lev Vygotsky are the most influential figures in constructivist theory. Piaget's cognitive constructivism emphasizes individual cognitive development, where learners build

understanding through processes of assimilation and accommodation. Knowledge is not given; it is discovered and reconstructed through interaction with the environment (Piaget, 1970).

Jerome Bruner further developed constructivist ideas through his concept of **discovery learning**, suggesting that learners should be encouraged to discover principles for themselves, guided by teachers who provide scaffolding. His notion of the *spiral curriculum* applies effectively to language learning, especially pronunciation practice, where learners refine earlier knowledge through iterative exposure (Bruner, 2009). In the EFL classroom, constructivist approaches promote learner autonomy, critical thinking, and contextualized communication. Multimedia tools embody these principles by offering authentic, interactive environments which allow learners to construct knowledge through experience. Constructivism thus serves as a foundation for technology-enhanced language learning, aligning with modern pedagogical shifts toward student-centred instruction.

Figure 1.6

Bruner's spiral curriculum



Note. Adapted from Rose, 2008, p. 161

The idea of a spiral curriculum came from Bruner (1960, p. 35). He asserted that "Any subject can be taught effectively in some intellectually honest form to any child at any stage of development … and we can come back to the same topic again and again, each time at a higher level." Instead of teaching something once and finishing it, we teach it many times, but each time:

- A little deeper,
- A little more complex,
- Connected to what the student already knows.

It looks like a spiral that we keep returning to the same ideas, but we rise higher every time.

2.1.6. Social Constructivism and Collaborative Learning

Social constructivism, grounded in the work of Lev Vygotsky (1978), extends constructivist theory by emphasizing the social nature of learning. Vygotsky argued that knowledge is co-constructed through social interaction and cultural mediation, highlighting the central role of language and collaboration. His concept of the Zone of Proximal Development (ZPD**)** remains foundational in understanding how interaction fosters growth. In EFL learning, social constructivism underlines the importance of dialogue, peer feedback, and collaborative tasks. Learners refine their pronunciation through corrective feedback, modelling, and shared reflection during group activities or online discussions. Digital platforms now extend this collaboration across boundaries, enabling learners to engage with peers and native speakers through social media, virtual classrooms, and voice-based applications.

Social constructivism is the idea that we learn best together, not alone. Thus, we build knowledge in social groups by talking, listening, and working with other people in the framework of three principles:

1. Learning is social first, individual second.
2. Zone of Proximal Development (ZPD) as a space between what I can do alone and what I can do with help from a friend or teacher. The best learning environment happens inside this zone with "scaffolding" as help slowly removed.
3. Collaborative learning is the classroom version of social constructivism (Johnson & Johnson, 1999). Students work in pairs or small groups, help each other, and explain things that they learn more than when the teacher explains alone.

Collaborative learning, a pedagogical application of social constructivism, involves learners working together to solve problems, complete projects, or discuss linguistic concepts. Johnson and Johnson (1999) identify positive interdependence, individual accountability, and promotive interaction as key features of effective collaboration. When integrated with multimedia technologies collaborative learning enhances both motivation and accuracy. Thus, social constructivism and collaborative learning place interaction and communication at the core of EFL pedagogy. They align perfectly with the multimedia classroom in the 21st century, where learning occurs through shared meaning-making, negotiation, and community engagement.

Furthermore, Connectivism, a relatively recent theoretical development proposed by Siemens (2005), extends these classical frameworks by emphasizing learning as an interrelated process, which uses multi-media and digital technologies. In EFL contexts, learners practice pronunciation skills not only through teacher feedback but also through exposure to online communities from all around the world, through recorded videos and podcasts, and speech recognition applications.

2.2. Educational Psychology and Learning Theories

Learning is a pivotal concept in psychology, which is the process of mental processes that occur in the mind of the learner. It has been investigated and probed by many scientists, thinkers, and educationalists for centuries since the birth of human civilizations until today. According to the historical chronicles and the human civilizations' scripts, the process of learning goes back to the beginnings of human civilizations and the invention of writing then it was inherited and developed over generations to the present time. In educational psychology, which is a branch of psychology that involves the study of the processes of learning and teaching languages and other disciplines. As a starting point, for many linguists, the source of the human language was divine which is the most credible hypothesis in historical linguistics that traces the origins of language. Yet, many assumptions have been postulated but they have never been approved and the artificial language like Esperanto showed the deficiency of the unnatural languages to survive in communities. Academically speaking, educational psychology lies at the centre of understanding how learners acquire, process, and retain knowledge. It provides the theoretical foundation on which effective teaching strategies and learning environments are based. This discipline specifically examines how cognitive, emotional, and social factors influence the learning process, borrowing from both psychology and pedagogy. In the 21st century, rapid technological advancement has been constantly reshaping how educational psychologists perceive learning, shifting focus from teacher-centred instruction to learner-centred and technology-mediated environments (Schunk, 2020).

Learning theories, traditionally framed within behaviourism, cognitivism, and constructivism, have evolved to accommodate the demands of contemporary digital learning. Behaviourism, for instance, proposed by B.F. Skinner (1953) and others, emphasizes observable

changes in behaviour as evidence of learning. Reinforcement and repetition techniques remain useful in language learning, particularly in pronunciation drills and multimedia-based feedback systems. Cognitivism, influenced by theorists like Jean Piaget and Jerome Bruner, views learning as an internal process involving memory, perception, and problem-solving. In the context of multimedia learning, cognitive theories are used to help explain how learners integrate visual and auditory information for improving pronunciation (Mayer, 2009).

Within educational psychology, learning theories are the frameworks that describe how learners acquire, process, and retain knowledge or skills. In the previous learning theories psychological science introduced many explanations to improves learning, then linked them to major learners.

The intersection of educational psychology and learning theories, hence, highlights the dynamic nature of learning in the 21st century. Understanding these frameworks is crucial for examining how embedding multimedia tools can thoroughly enhance EFL learners' pronunciation, as they provide genuine insights into the cognitive and behavioural mechanisms that support language acquisition in settings supported by technology. Schunk (2020) remarks in this respect: "The explosion in technology that has occurred in recent years has yielded new methods that can show how the brain functions while performing mental operations involving learning and memory. The data yielded by these new methods are highly relevant to classroom teaching and learning and suggest implications for learning, motivation, and development" (p. 30).

2.3. Multimedia Integration in Language Pedagogy

The integration of multimedia in language pedagogy represents one of the most significant transformations in educational practice in the 21st century. Multimedia refers to the combination of text, sound, images, video, and animation to create interactive and multisensory learning

environments. In language education, such practice has evolved from only providing visual aids as supplementary to become a central pedagogical component necessary in language acquisition. As in Mayer's CTML (2009), meaningful learning occurs when learners actively integrate verbal and visual information within their cognitive structures, promoting deeper understanding and longer memorizing.

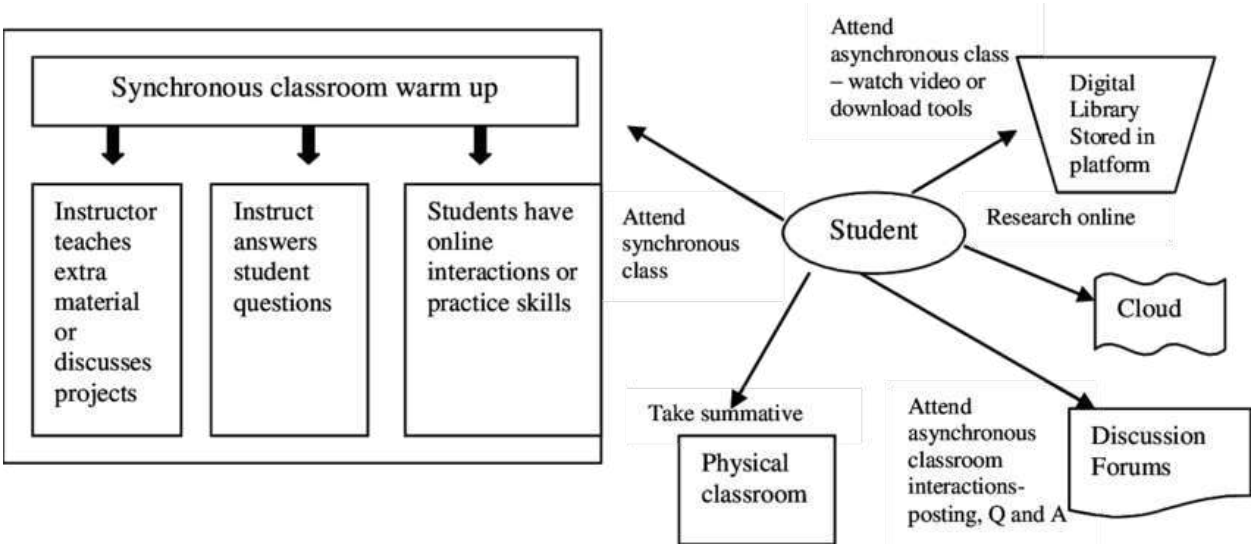
From a theoretical point of view, multimedia integration is rooted in Constructivist and Sociocultural learning theories, which view knowledge as actively constructed through experience, interaction, and reflection. Multimedia tools, such as language learning applications, interactive pronunciation software programs, and videos, provide authentic contexts that engage multiple senses and different learning styles. According to Chapelle (2003), computer-assisted language learning (CALL) environments help learners experimenting with language, receive immediate feedback, and monitor their progress, aligning with Vygotsky's concept of the *Zone of Proximal Development* (ZPD). These interactive tools thus serve both as scaffolds and mediators of linguistic and cognitive development.

The multimodal feedback mechanism also enhances awareness of phonetic detail and supports self-correction, key aspects of autonomous learning. As Chun (2012) notes, digital environments that merge audio-visual tools help learners build stronger form/meaning connections, promoting more accurate speech. Moreover, multimedia integration promotes learner-centred pedagogy by accommodating diverse learning preferences and encouraging self-paced exploration. It transforms the teacher's role from sole knowledge provider to facilitator and guide. This aligns with Blended Learning and Flipped Classroom models, where multimedia materials extend learning beyond the classroom and foster continuous engagement.

Using multiple modes supports comprehension and memory by presenting language in diversified, contextualized ways that attract the different senses and learning preferences. When grounded in learning theories such as connectivism (such as multimedia learning principles and communicative language teaching), multimedia can increase motivation, provide authentic input, and create opportunities for meaningful interaction and autonomy. Moreover, many practical applications such as using language labs, CALL platforms, videos, simulations, or mobile apps could develop listening, speaking, reading, and writing skills in the EFL context.

Figure 1.7

Holistic Flipped Classroom Model



Note. Adapted from Ozdamli & Asiksoy, 2016, p. 103

Multimedia integration in language pedagogy is not merely a technological innovation but a theoretical advancement grounded in cognitive and constructivist principles. It enhances interactivity, promotes deeper cognitive processing, and creates authentic, personalized learning experiences. As digital literacy becomes essential in education, multimedia stands as a bridge between traditional language teaching and the dynamic, interactive learning environments necessary for learners in the 21st century.

2.3.1. Dual Coding Theory and Multimodality

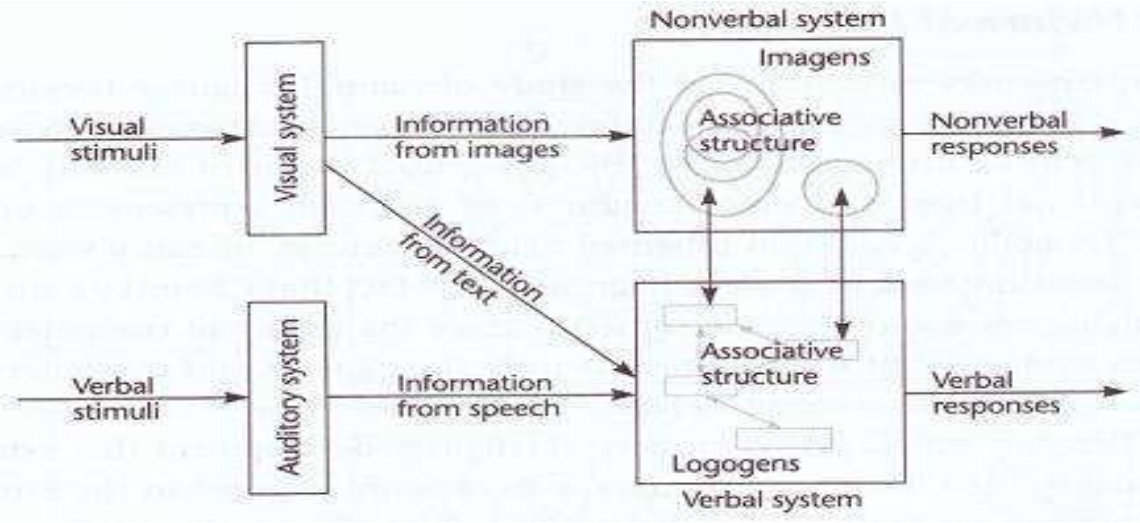
The integration of multimedia in TEFL (Teaching English as a Foreign Language) finds strong support in Dual Coding Theory (DCT) and the broader framework of multimodality. Proposed by Allan Paivio (1971), DCT suggests that human cognition processes information through two distinct but interconnected channels:

1- The verbal system, this is the collection of linguistic input such as written and spoken words,

2- The non-verbal system, which includes imagery and visual representations. Learning becomes more effective when both channels are activated simultaneously.

Figure 1.8

Paivio's Dual-Coding Theory



Note. Adapted from Paivio & Clark, 2006, p. 166

In the context of language teaching, DCT underlines the pedagogical value of visual aids, videos, and interactive multimedia that complement verbal input. For instance, pairing vocabulary items with corresponding images, or using video clips to demonstrate cultural contexts, helps

learners create stronger mental associations. This multimodal processing supports deeper learning, particularly for EFL learners who benefit from contextual and visual reinforcement.

Multimodality, as discussed by Kress and van Leeuwen (2001), extends this principle by recognizing that meaning is constructed through the interaction of multiple modes: visual, auditory, textual, and kinaesthetic. Language learning in multimedia classrooms becomes an active process of interpreting and combining these modes. Teachers who design lessons incorporating sound, text, image, and movement actually adopt *multiliteracies* that go beyond linguistic competence, preparing learners for the complex communicative demands of today's digital age. Dual Coding Theory and Multimodality jointly emphasize that the human mind learns best when information is processed across different senses. Their application in TEFL encourages a holistic and engaging approach to language teaching, making multimedia not just as a supplementary but as a cognitive tool for better understanding.

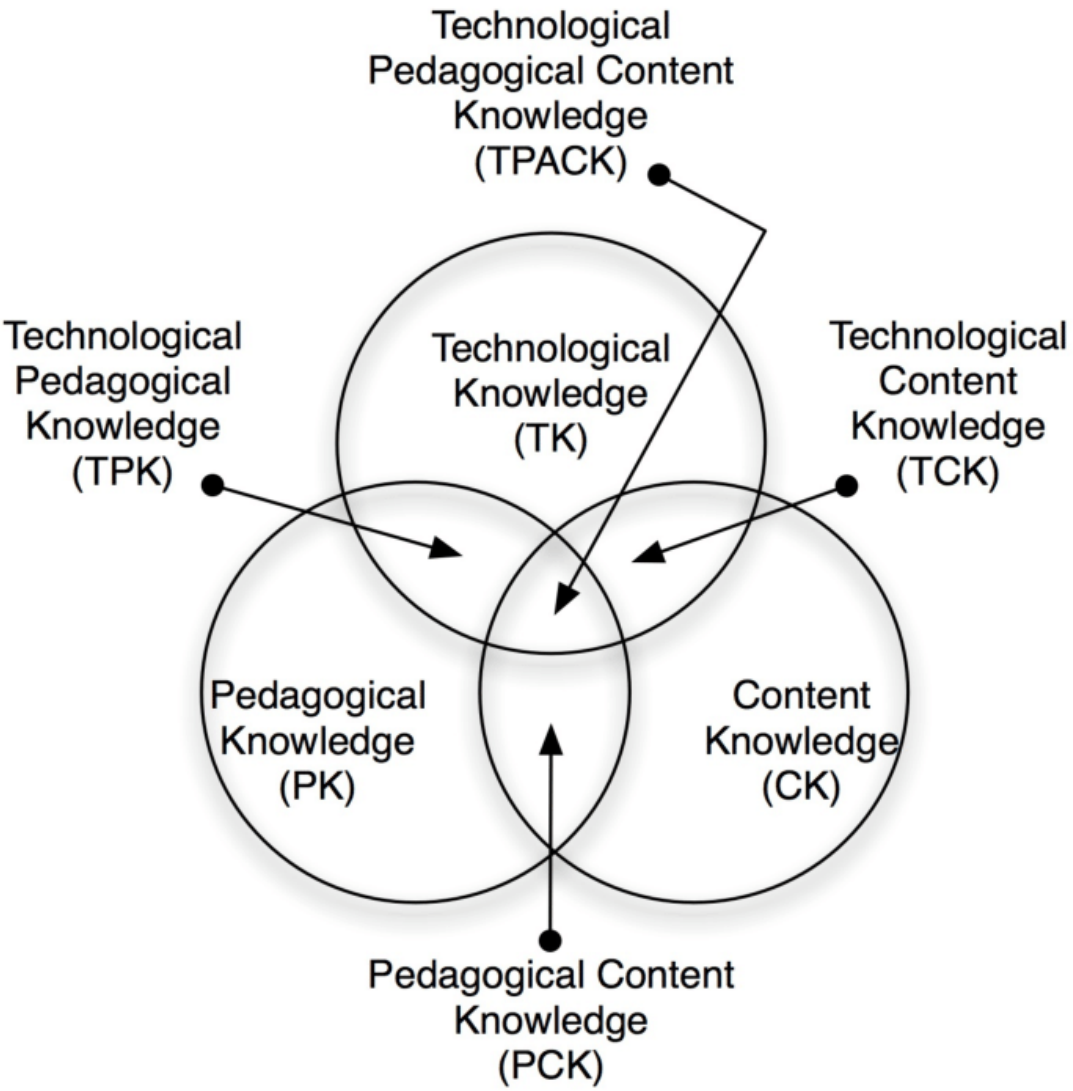
2.3.2. Teachers' Professional Development in Multimedia Literacy

The effective integration of multimedia in TEFL largely depends on teachers' digital competence and multimedia literacy. In modern educational contexts, teachers are no longer mere transmitters of knowledge but designers of learning environments who must navigate and orchestrate multiple modes of representation. Developing multimedia literacy is essential for teachers to exploit the full potential of digital tools in enhancing language learning. According to Mishra and Koehler's (2006) TPACK model (Technological Pedagogical Content Knowledge), effective technology integration stems from the intersection of three knowledge domains: content, pedagogy, and technology. EFL teachers must not only understand language pedagogy but also develop the ability to select, adapt, and implement digital media that align with specific learning objectives.

For Medgyes (2017, p. 131), the teacher education and training for non-native speakers is restricted to the institution in the pre- and in-service training as a part of teacher development. Nevertheless, it goes beyond to include any voluntary activity pursue with the intention of enhancing his/her professional expertise.

Figure 1.9

Technological Pedagogical Content Knowledge



Note. Adapted from Jamieson-Proctor, 2012, p. 27

Multimedia literacy entails understanding how meaning is shaped through multimodal resources. As Jenkins (2009) argues, educators must develop new literacies that enable both teachers and learners to participate actively in digital culture through creation, collaboration, and communication. This approach empowers teachers to guide students in interpreting and producing multimodal texts, from podcasts and graphs to digital projects. Ultimately, continuous professional development in multimedia literacy encourages teacher autonomy, creativity, and innovation. It encourages educators to view technology not as an *add-on* but as an integrated component of communicative and interactive language teaching.

2.3.3. Multimedia and Oral Proficiency

The integration of multimedia into English as a Foreign Language teaching has revolutionized the development of oral proficiency, which encompasses learners' ability to communicate effectively, fluently, and accurately in English. Traditional classroom instruction often relied on limited exposure to authentic speech and repetition drills. However, multimedia classrooms provide learners with rich linguistic input and realistic dialogues. By combining sound, text, video, and other interactive outlets, multimedia enhances the cognitive aspects of language learning (Chapelle, 2009). From a cognitive perspective, multimedia offers dual-channel processing, where learners simultaneously receive auditory and visual stimuli, hence facilitating deeper understanding of language patterns (Mayer, 2014). The visual representation of spoken language through gestures, facial expressions, and situational cues helps understanding meaning and perfecting pronunciation. Learners can observe and imitate authentic models, improving their rhythm, stress, and intonation patterns.

Ultimately, the integration of multimedia into oral language instruction promotes communicative competence, bridging the gap between controlled classroom practice and real-life

communication. Through multimedia learning, learners are not only exposed to the sounds and structures of English but also to the cultural and contextual nuances that shape authentic conversation.

2.3.4. The Role of Multimedia in Enhancing Speaking Fluency

Multimedia tools, such as videos, podcasts, speech recognition software, and interactive applications play a crucial role in promoting speaking fluency in EFL classrooms. Unlike traditional methods, multimedia provides learners with authentic input and multimodal contexts that simulate real-life communication. Through repeated exposure to native speech, visual cues, and contextualized dialogues, learners develop a better sense of rhythm, intonation, and stress patterns (Gilakjani, 2012).

Moreover, multimedia encourages autonomous and interactive learning, allowing students to engage in tasks that promote spontaneous speech. Technologies such as AI-assisted pronunciation applications or online speaking platforms provide immediate feedback, which is critical for improving fluency and reducing fossilized errors. Studies show that integrating video-based speaking tasks improves learners' ability to organize thoughts quickly and communicate more naturally. The visual dimension of multimedia enhances cognitive processing and encourages dual coding and the integration of verbal and visual information which helps learners store and retrieve linguistic information more efficiently (Paivio, 1990). Thus, multimedia not only exposes learners to rich input but also reinforces the psycholinguistic mechanisms underlying fluent speech production.

2.3.5. Authentic Materials as Multimodal Resources

Authentic materials such as interviews, documentaries, films, and YouTube content serve as multimodal resources that expose learners to real-life language use. They combine verbal, visual, and contextual cues, helping learners interpret meaning beyond words. According to Guichon and McLornan (2008), the use of multimodal input enhances learners' pragmatic and sociolinguistic competence by linking form, meaning, and context. In EFL classrooms, authentic multimedia materials support situated learning, allowing learners to see how language functions in social interactions. Visual and auditory elements reinforce comprehension and pronunciation while offering models of natural discourse. This multimodality also caters to different learning styles: visual, auditory, and kinaesthetic (Mayer, 2014).

Authentic materials are meant to be used in real contexts, such as newspapers, menus, podcasts, films, social media posts, or websites representing the target language use. Accordingly, EFL learners are exposed to natural language use, current topics, and cultural norms, which can increase motivation and support real-life communication skills. Particularly, using these materials in several modes, written text, images, layout, sound, and sometimes movement or interaction that learners interpret together. For example, a news video that offers spoken language, on-screen text, visuals, and possibly music, giving multiple complementary cues that strengthen comprehension and retention.

In language teaching, authentic multimodal materials are used to design tasks where learners read, listen, view, and respond in integrated ways, mirroring how language is used outside the classroom. Teachers may scaffold these materials with pre-teaching vocabulary, guiding questions, and follow-up speaking or writing activities so learners can benefit from the richness without being overwhelmed.

2.3.6. Challenges in Implementing Multimedia-Based Learning

While multimedia-based instruction has transformed language education, its successful implementation is still limited by a wide range of pedagogical, technical, and institutional challenges. These challenges often determine whether technology serves as an enhancement to learning or becomes an obstacle to effective teaching. One major issue is teachers' digital competence. Many EFL teachers lack sufficient training in selecting, adapting, and integrating multimedia resources effectively into their lessons (Hampel & Stickler, 2015). The absence of systematic professional development in multimedia pedagogy leads to superficial use of digital tools without meaningful interaction or learner autonomy. Developing technological pedagogical content knowledge is still a necessary for ensuring that multimedia use is grounded in sound teaching principles.

Infrastructure and accessibility also pose significant problem, especially in developing contexts where internet connectivity, hardware availability, and maintenance are inconsistent. These limitations can reduce students' exposure to multimedia tasks and create inequities in access to digital learning opportunities (Reinders & Benson, 2017). Without institutional investment in reliable technology and technical support, even the most innovative multimedia approaches risk remaining theoretical rather than practical. Additionally, teachers face challenges in material evaluation and cultural appropriateness. Many multimedia resources are designed for global audiences, which may not align with local cultural norms or linguistic varieties. Adapting such materials to suit learners' sociocultural backgrounds requires both critical awareness and pedagogical creativity.

2.4. Pronunciation in Language Learning and Teaching

To start with, the forthcoming section reviews the literature concerning the aforementioned foundation on educational psychology and learning theories to end with the latest trends methods and strategies of teaching pronunciation and the best ways of applying them. More specifically, pronunciation is an essential aspect of language proficiency, mainly, it affects learners' communicative abilities and competencies. As pointed by Celce-Murcia, Brinton, and Goodwin (2010), “pronunciation is not an optional extra for the language learner, but an integral part of the language system” (p. 8). In recent years with rise of social media platforms, pronunciation pedagogy gained more research and attention from learners, education stakeholders and pedagogists.

In recent years with rise of social media platforms, pronunciation pedagogy gained more research and attention from learners, education stakeholders and pedagogists. Pronunciation is not an extra skill but it is the cornerstone of speaking. For (Celce-Murcia et al., 2010), bad pronunciation is a source of unintelligibility and comprehensibility where the listener will not understand even perfect grammar and vocabulary in speech. In their book, they reviewed the history of pronunciation teaching which was named as the Cinderella of speaking, that is after a long focus on pronunciation during the era of audiolingual paradigm then it was totally overlooked with the rise of direct method then afterwards.

Table 2.2

The History of Pronunciation Pedagogy

PERIOD	TASKS AND FOCUS	PROBLEMS
1950S–1960S (ALM)	Lots of drills, perfect accent was the only goal	Students sounded robotic, gave up when impossible
1970S–1980S	Almost no pronunciation – focus on grammar	Students could write but not speak
1990S–2000S	"Teach only what is necessary for being understood"	Better, but many teachers still ignored it
2010–2025	Goal: **intelligibility & comprehensibility**	(Celce-Murcia, 2010; Saito, 2012; Derwing, & Munro, 2015; Levis, 2020) This study belongs to this period of focus.

2.4.1. Foundational Concepts in Pronunciation Pedagogy

In pronunciation pedagogy, prioritizing communication with a special reference to the components of fluency intelligibility, comprehensibility, and accentedness. In recent years, pedagogy has shifted from targeting native-like accuracy to a communicative orientation prioritizing intelligibility and comprehensibility. In the same regard, Munro and Derwing (1995) established the tripartite model based on accentedness which is the difference between someone's speech and a reference accent in addition to the intelligibility of the listeners' clarity of spoken language to understand as well as comprehensibility which is the perceived ease or difficulty with which a listener understands speech. On the other hand, Levis (2005) differentiates "Intelligibility

Principle" and "Nativeness Principle," asserting that "trying to become nativelike is almost never the appropriate goal for L2 learners because it is nearly impossible and unnecessary for communication" (p. 370). Instead, emphasis should be placed on features that impede understanding, as he adds (Levis, 2005, p. 369) "The history of pronunciation in English language teaching is a study in extremes. Some approaches to teaching, such as the reformed method and audiolingualism, elevated pronunciation to a pinnacle of importance, while other approaches, such as the cognitive movement and early communicative language teaching, mostly ignored pronunciation".

Determining which pronunciation features to teach involves the concept of "functional load," reflecting how much work a phoneme does to distinguish word meanings. Catford (1987) and Brown (1988) reckoned this, showing sounds like /l/-/r/ in English distinguish many minimal pairs and thus carry a high functional load. Munro and Derwing (2006) demonstrated that errors affecting high functional load pairs "significantly impact comprehensibility and accentedness more than low functional load errors". Teaching should thus prioritize high functional load errors, as "emphasis on high FL errors should lead to greater improvement in how L2 speakers are understood" (Munro & Derwing, 2006, as cited in Levis, 2005, p. 371).

For teaching suprasegmentals in English like stress, rhythm, and intonation, a major shift in pronunciation pedagogy was the recognition of suprasegmentals—as central to comprehensibility. Accordingly, for the Arab learners of English, Benrabah (1997) pointed out that misplaced word stress, among Algerian EFL learners, was a source of unintelligibility in English. Prominence in stress is essential for intelligibility for L2 listeners, with Jenkins (2000) noting its role in international English communication. More specifically, phonetic training was a predominant instructional approach that incorporates explicit phonetic training and the use of

imitation or mimicry. According to Celce-Murcia (2001), “the most successful methodologies in L2 teaching and learning have included repetition and imitation of words and sentences to a greater or lesser extent”. Flege (1988) also demonstrated that “imitating native speakers could help learners acquire the natural intonation and rhythm of a language, leading to more fluent and comprehensible speech”. In modern pedagogy, learners study the International Phonetic Alphabet (IPA), “teaching them to recognize and produce the distinct sounds of their target language, as well as practice tongue placement, lip rounding, and other articulatory features” (Derwing & Munro, 2015).

From another regard, minimal pairs are used in teaching English as a method in pronunciation teaching. For example, “ship” and “sheep” differ in only a single phoneme. Hansen (1995) found that “language teachers can improve their students’ pronunciation markedly drilling minimal pairs in order to help them improve their intelligibility,” Meanwhile, spelling can both support and hinder pronunciation learning. English is opaque when it comes to orthography often complicates prediction of pronunciation from spelling (see Chomsky & Halle, 1968). Another important point, addressing variation in English dialectal differences pronunciation varies in the English-speaking countries. So, teachers must be aware of both “positional variation” (phonemes change depending on where they occur in a word) and “dialectal variation” (regional accents). Levis (1999) asserts that “teachers should be encouraged to model their own dialect but should allow students to use any acceptable variation in pronunciation,” fostering realistic and inclusive pronunciation standards. Nowadays, in the digital age of technology and pronunciation, recent trends emphasise the significance of integrating multimedia tools and digital materials in pronunciation pedagogy. Several platforms provide “authentic speech models, automated

pronunciation grading, and real-time feedback". As pointed by Sanako (2024), "technology tools can play a vital role in building students' English or Foreign language pronunciation skills."

2.4.2. Teaching and Learning Phonetics and Phonology

Phonetics and phonology represent two core subfields of descriptive and theoretical linguistics as a window of understanding pronunciation in language learning and teaching. Despite considering them interchangeably, their discrepancies are both theoretically and pedagogically fundamental in the process of pronunciation teaching. As Delahunty and Garvey (2010) assert, "we begin with phonetics, a system for describing and recording the sounds of language objectively," while "phonology concerns itself with the ways in which languages make use of sounds to distinguish words from each other" (p. 90). A well-defined distinction is needed to help EFL educators facilitate pronunciation pedagogy with both practical drills and phonological understanding. Scientifically speaking, phonetics is the scientific study of human spoken sounds in terms of their physical properties, without considering their linguistic function. It focuses on how sounds are produced using speech organs (articulatory phonetics), their acoustic properties of the physical features of the produced soundwaves (acoustic phonetics), and how they are perceived and perceived from the part of the listener (auditory phonetics). As Roach (2009) defines it, "Phonetics is the scientific study of speech sounds. It focuses on the physical properties of sounds and how they are produced, transmitted, and perceived" (p. 1).

In the educational context, phonetics allows specialists to record, categorise, and analyse speech by developing "a system for classifying speech sounds on the basis of how they are produced," most notably using the International Phonetic Alphabet (IPA), where "every phonetic symbol represents an actual sound" and "no 'silent' letters" exist (Delahunty & Garvey, 2010, p.

90). For example, the phoneme [p] in *pat* and the phoneme [b] in *bat* are both bilabial stops, differentiated mainly by voicing and force of articulation; such properties are accurately described in phonetic terms to improve EFL learners' pronunciation segmental features.

The table below illustrates the vocal apparatus and key articulators involved in speech production, as adapted from Delahunty & Garvey (2010):

Table 2.3

Vocal apparatus and key articulators

Articulator	**Description**
Vocal folds	Vibrate to produce voicing
Tongue	Key to shaping sounds, placement varies
Lips	Create bilabial sounds, rounding, etc.
Velum (soft palate)	Controls oral/nasal airflow

Phonology, in contrast, investigates how the sound system of a particular language functions within the sound pattern of that language. It describes and explains the abstract, cognitive aspects of the sound system as distinct sound units—the phonemes, their distribution, and the patterns that are ruled by phonotactics and the permissible sounds combinations and change across contexts. As Roach writes, "Phonology is the study of how sounds function within a particular language or languages. It deals with the abstract, cognitive aspects of sounds as linguistic units" (2009, p. 5). Correspondingly, the essential concepts in phonology are the phonemes which are defined as "The smallest unit of sound that can distinguish meaning in a language" (Roach, 2009, p. 8). In addition, the phonemic inventory represented in the set of distinct phonemes in a language

as well as the p**honological rules of the** patterns governing permissible combinations and contextual alternations within the sound system of the language. In the same vein, it analyses why, for example, the English [p] and [b] contrast (as in *pat* vs. *bat*), but two allophones of [p] (such as [p^h] and [p]) do not create meaning differences.

The interaction between phonetics and phonology can be approached as the distinction between the *physical description of phonemes* and its *functions, abstract organization of the sound combinations* within the particular language system. First, phonetics analyses the different physical realisations or "phones" of sounds, e.g., any sound articulated by native speakers such as [p] is slightly different phonetically yet is considered as the same classification. Whereas phonology organizes these physical varieties into functional categories (phonemes and allophones), focusing on the sound contrasts that signal meaning "while phonetics deals with the physical and measurable aspects of sounds, phonology focuses on the functional and abstract aspects of sounds as part of a language system" (Roach, 2009, p. 11). To illustrate more, in English, [p^h] (aspirated as in *pin*) and [p] (unaspirated as in *spin*) are both allophones of the phoneme /p/, not meaning-distinguishing. But /p/ and /b/ contrast in meaning; they are separate phonemes.

Table 1.4

The hierarchy of "Phones, Allophones, Phonemes"

Level	**Example**	**Linguistic Function**
Phone	[p^h], [p]	Physical realization
Allophone	[p^h], [p] variants	Contextual variant, not contrastive
Phoneme	/p/, /b/	Meaningful contrast (minimal pairs)

Note. Adapted from Delahunty & Garvey, 2010.

As an investigation of the significance in pronunciation for understanding both phonetics and phonology is essential for EFL instructors, Delahunty and Garvey stated that "Phonetics provides the tools to describe and measure sounds, while phonology explains how those sounds are used and patterned in language" (2010, p. 90). For learners, phonetic exercises can improve perceptual accuracy, while phonological insight can clarify which distinctions matter for communication in English. Roach (2009) emphasizes the interconnectedness of the two fields, noting, "comprehensive study and analysis of spoken language cannot be achieved without recourse to both phonetics and phonology" (p. 16). He further affirms, "Knowing the difference between physical sound and its function in a language is crucial, especially in teaching pronunciation to non-native speakers, where both sound and meaning matter" (Roach, 2009, p. 18). Defining a clear-cut distinction between phonetics and phonology offers a grounded basis for an effective pronunciation instruction. For Algerian EFL settings, there is an explicit training and illustration on both segmental and prosodic features of English pronunciation to improve EFL learners' pronunciation and enhance pedagogical approaches in teaching English sounds and suprasegmentals.

2.4.3. Segmental and Suprasegmental Features in Fluency Development

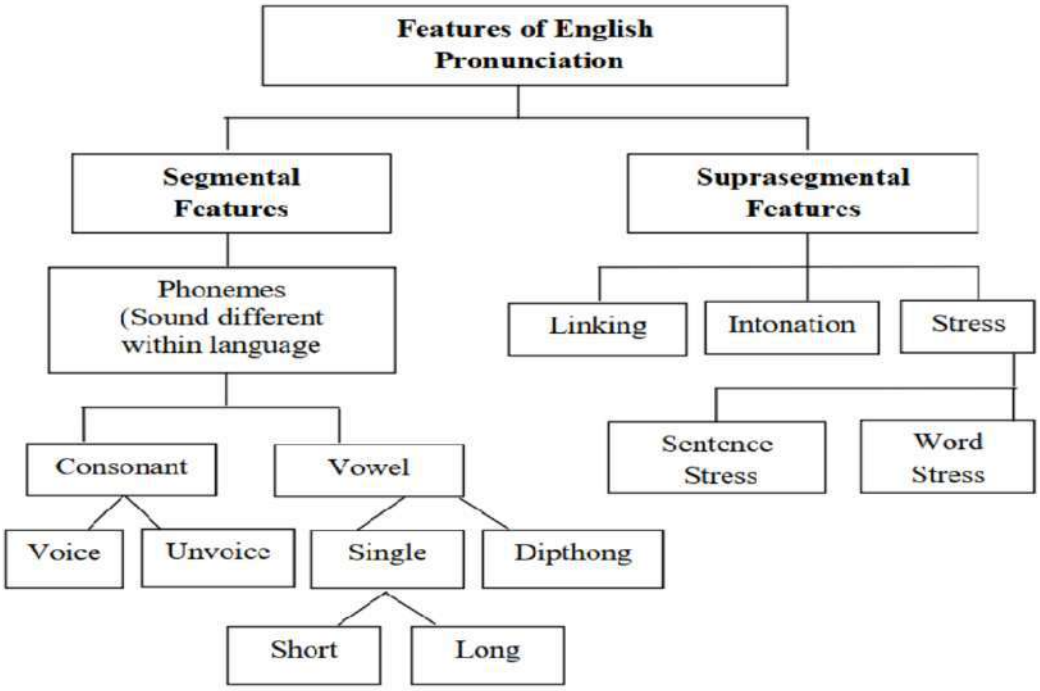
In pronunciation, there is a dual focus on segmental and suprasegmental features, both of which are critical for developing intelligible and comprehensible with fluency development among the EFL learners. Understanding the theoretical foundations of the latter features could be used in fostering fluency to come up with pedagogical implications. Segmental features refer to individual phonemes either vowels or consonants, which constitute the smallest units of sound in a language.

Mastery of segmentals is often the starting point for pronunciation building in the EFL instruction, any failure or mispronunciation at this level can decrease intelligibility. Celce-Murcia et al. (2010) emphasize that "the misproduction of even a single phoneme can lead to communication breakdowns, particularly in contexts where minimal pairs (e.g., *ship* vs. *sheep*) are critical to meaning" (p. 45).

For instance, Arabic-speaking learners of English may struggle with the /p/ vs. /b/ or /n/ vs /ŋ/ distinction due to the absence of /p/ and /ŋ/ in their native phonological inventory. However, during the golden age of teaching pronunciation through using audiolingual method, there was an overemphasis on segmentals has been subjected to major criticism. Derwing and Munro (2005) argue that "while segmental accuracy is important, exclusive focus on individual sounds risks neglecting the broader prosodic features that underpin natural speech" (p. 385). This critique aligns with the idea that segmental errors in pronunciation, while noticeable among EFL learners, are often less troublesome to overall intelligibility than suprasegmental inaccuracies (Field, 2005).

Figure 1.10

Features of English pronunciation



Note. Adapted from (Pourhosein, 2012)

Suprasegmental features encompass the rhythm of communication represented in the prosodic stress, rhythm, intonation, and connected speech aspects. These elements govern the flow and melody of language, shaping how utterances are perceived emotionally and pragmatically. Gilbert (2008) compares suprasegmentals to the "musical score of speech," asserting that "without attention to prosody, even phonetically accurate speech may sound robotic or emotionally flat" (p. 72).

1. **Stress and Rhythm**: English is a stress-timed language, where stressed syllables occur at regular intervals, creating a rhythmic pattern. Misplaced stress can alter meaning (e.g., *REcord* vs. *reCORD*) or confuse listeners. Dalton and Seidlhofer (1994) note that

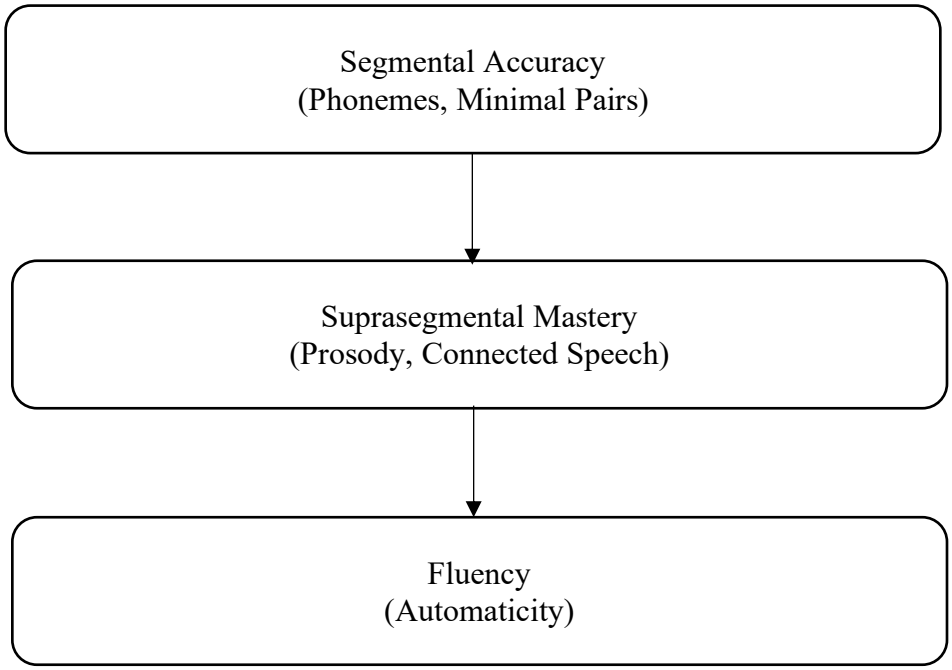
"learners who apply syllable-timed rhythms (common in Arabic and French) to English often produce speech perceived as monotonous or disjointed" (p. 89).

2. **Intonation**: Rising and falling pitch contours signal questions, statements, or attitudes. Brazil (1997) highlights that intonation errors may lead to pragmatic misunderstandings, such as interpreting a declarative sentence as sarcastic due to inappropriate pitch movement.
3. **Connected Speech**: Features like linking, elision, and assimilation (e.g., "gonna" for "going to") are essential for fluency. Brown (1990) argues that "failure to assimilate these features results in staccato-like speech, hindering natural interaction" (p. 112).

Fluency is defined as "the ability to produce language smoothly, rapidly, and accurately" (Lennon, 1990, p. 391). It relies on the incorporation of both segmental and suprasegmental competencies among EFL learners. Nation and Newton (2009) propose a hierarchical model

Figure 1.11

Hierarchy of Pronunciation Competence



(Figure 1.11) where segmental accuracy forms the foundation, but suprasegmental mastery enables higher-order fluency including speech rate and smoothness.

The model above indicates that while segmental training and teaching of phonemes fosters accuracy in pronunciation, while mastery of the suprasegmental features promotes fluency that is the ability to use language utterances in rapid speech without considerable effort. Jenkins (2000) approves this view, stating that "prosodic features act as 'glue,' binding discrete sounds into coherent utterances" (p. 159). For example, a learner may articulate /θ/ correctly in isolation (segmental) but struggle to embed it in rapid speech features with appropriate aspect of casual speech (suprasegmental) in the connected speech. It is recommended to:

1. **Balanced Instruction**: Effective pedagogy must integrate both features. Derwing et al. (1998) advocate for a "top-down" approach, where suprasegmental practice precedes segmental refinement in early stages to build communicative confidence.
2. **Contextualized Practice**: Activities like role-plays or discourse analysis help learners apply prosodic features in authentic contexts. Chela-Flores (2001) recommends using dialogues with marked intonation patterns to teach pragmatic functions (e.g., shadowing, politeness).
3. **Technology Integration**: Multimedia tools, such as waveform visualizations or pitch trackers, which can make abstract suprasegmental features tangible (Levis, 2007). For instance, softwares like ASR and CAPT or Praat allow learners to compare their intonation patterns with native speaker models.

The dichotomy between segmental and suprasegmental features is a false one; both are indispensable for pronunciation pedagogy and fluency development. As Munro and Derwing

(2011) conclude, "The goal is not to choose between sounds and prosody but to recognize their symbiotic relationship in shaping intelligible, fluent speech" (p. 201). In the future sections of the study, we will explore how embedded multimedia can bridge this division by providing multimodal input and interactive practice opportunities as stated by (Saito, 2018), "late L2 learners can refine pronunciation of segmentals through exposure and focused practice, regardless of lexical context". Classically, Wells (1996) argued that "teaching segmental features is the foundation of pronunciation: the learner must master the inventory of sounds and their articulatory properties before moving on to higher-level features" (p. 74).

According to Lehiste (1970), "prosodic aspects of language are superimposed on its segmental, syllabic, and lexical features and are, hence, regarded as suprasegmental," constituting properties such as syllable structure, stress, and pitch variation. These elements work in concert to provide speech with its distinctive melody and affect.

Table 2

Main Suprasegmental Features in Speech

Feature	Function	Example
Stress	Highlights key syllables	REcord (noun) vs. reCORD (verb)
Rhythm	Regulates timing	"The cat sat on the mat"
Intonation	Modulates pitch contour	"You're leaving?" (rising)
Pausing	Marks phrase boundaries	"Let's eat, Grandma"

Note. Adapted from Grant, 2023

In recent research, where emphasis was laid more on fluency over accuracy and imitation, there is an increasingly focus on suprasegmentals often have a greater impact on listener understanding than individual segmental errors (Anderson-Hsieh et al., 1992). As Derwing et al. (1998) emphasize, "suprasegmentals play a more important role than segmentals (phonemes) in the intelligibility of speech pronunciation and comprehensibility in communicative contexts." Hahn (2004) found that "predictable prominence patterns of international teaching assistants resulted in listeners remembering more information, while misplaced prominence impaired intelligibility." In this regard, in the late 1960s there was a pedagogical shift, with teachers and researchers advocating for a stronger focus on suprasegmentals in L2 instruction. According to Murphy (2010), "prosody represents the speaker's management of the speech stream and is crucial for listener perception and meaning-making."

2.4.4. Pronunciation and Language Skills Integration

Pronunciation is no longer viewed as an isolated linguistic component but as an integral part of overall communicative competence. Effective language instruction recognizes that pronunciation interacts dynamically with other skills, particularly listening and speaking, to shape intelligible communication. The integration of pronunciation into broader language instruction ensures that learners develop both perceptual and productive phonological awareness.

Pronunciation, is not a separate skill, it is an integral part that holds all four skills together. In speaking the first barrier that many foreign language learners face in mispronunciation, which affects their listening, speaking, and even reading. Pronunciation is the bridge between all the four language skills. Teaching it with multimedia materials makes listening easier, speaking clearer,

reading more natural, and vocabulary memorised longer. It can be enhanced through practice as in the following table:

Table 2.6

The Integration of Pronunciation in Teaching Language Skills

LANGUAGE SKILL	PRONUNCIATION IMPACT	PRONUNCIATION EFFECT	PRACTICAL SOLUTIONS
LISTENING	Good pronunciation better ear for sounds and stress	Students hear "ship" but think it is "sheep"	Minimal-pair listening Praat waveform Listening for ear training
SPEAKING	Clear pronunciation People understand you	Teacher asks "What's the time?" student says "tenk you" nobody understands	Shadowing real dialogues Corrective feedback Clear speech
READING ALOUD	Correct stress and intonation make reading natural	Reading like Arabic (every syllable same length)	BBC scripts record compare with native re-read

VOCABULARY LEARNING	Hearing the word correctly Remember it	Students forget when they are exposed to authentic content repeatedly	YouTube pronunciation Quiz
GRAMMAR IN SPEECH	Intonation shows questions, lists, emphasis	"Let's eat grandma" vs "Let's eat, grandma"	Intonation arrows on screen role-play practice

2.4.5. Interplay Between Pronunciation and Listening

Listening and pronunciation are interdependent processes; accurate perception of sounds forms the foundation for accurate production. According to Field (2008), learners must first be able to differentiate between phonemes before they can reproduce them effectively. Exposure to authentic audio materials such as dialogues, films, and podcasts, enhances learners' ability to recognize stress, rhythm, and intonation patterns, which subsequently improves pronunciation. In this regard, listening serves as a model and feedback mechanism, allowing learners to internalize target language features. Multimedia tools, especially those offering visual and auditory feedback, reinforce this interrelatedness between sound to articulation.

2.4.6. Pronunciation Role in Speaking Fluency

Speaking fluently relies heavily on accurate pronunciation, as it facilitates both speed and clarity. Gilakjani (2012) argues that effective pronunciation instruction enhances learners' confidence and communicative competence, leading to smoother speech delivery. In oral communication, segmental accuracy (individual sounds) and suprasegmental features (stress,

rhythm, and intonation) jointly determine fluency and comprehensibility. In EFL contexts like Algeria, where English exposure is limited, integrating pronunciation practice into speaking activities, such as role-plays, debates, and interactive software, helps learners move beyond accuracy toward naturalness in communication.

2.5. Debates in Pronunciation Standards

The field of pronunciation pedagogy is characterized by ongoing debates regarding which pronunciation norms teachers should promote. The rise of English as a global lingua franca has challenged the traditional emphasis on native-speaker norms, prompting a shift toward intelligibility and communicative effectiveness.

2.5.1. Native Speakerism versus Global Englishes

The idea of native-speakerism was very strong from the 1950s to the 1990s especially with the Audio-Lingual Method and early communicative teaching. Consequently, during that era most ESL and ESL learners felt ashamed of their accent and gave up speaking in public especially with natives. On the other hand, Global Englishes is the new reality (Jenkins, 2015; Seidlhofer, 2011), in the present time, non-native speakers make over (80 %) who use English more than native speakers. That is, English belongs to everyone who uses it named as World Englishes or English as a Lingua Franca. Hence, the real goal is mutual intelligibility, that is, being understood easily by any English speaker, not sounding British or American.

Historically, Received Pronunciation (RP) and General American (GA) served as models of correctness in EFL instruction. However, scholars such as Jenkins (2000) and Seidlhofer (2011) argue that this orientation toward native-speaker models perpetuates linguistic inequality and

disregards the global diversity of English users. The "Global Englishes" perspective promotes the idea that pronunciation should aim for mutual intelligibility among international speakers rather than mere imitation of native accents. This view aligns with the sociolinguistic reality of English use in multilingual settings like Algeria, where English functions primarily as a tool for international communication rather than social identity as represented in the table below:

Table 2.7

Comparison between Native-speakerism and Global Englishes

Idea	**Native Speakerism**	**Global Englishes (This Study)**
Model to Copy	Only BBC or CNN voice	RP or GA or any clear accent of English
Acceptable Accent	British RP or General American only	Any accent as long as the **Lingua Franca Core** is correct
Teacher's Accent	Must be native	Good if the teacher is intelligible and confident
Student's Goal	Zero accent	Clear, confident, proud of own voice
Example Sentence	"I want to speak like the Queen"	"I want people understand me easily in meetings"

2.5.2. Intelligibility, Comprehensibility, and Accentedness

Contemporary pronunciation research distinguishes between three related but distinct constructs: intelligibility (how well speech is understood), comprehensibility (how easily it is understood), and accentedness (how different it sounds from a native model) (Munro & Derwing, 1995). While native-like pronunciation often proves to be challenging for adult learners, achieving

intelligibility and comprehensibility remains a realistic pedagogical goal. Emphasizing these criteria shifts focus from eliminating accent to improving effective communication. In the Algerian EFL context, this approach encourages learners to develop clear and comprehensible pronunciation without the psychological burden of achieving native-like speech.

Table 2.8

The Difference between Intelligibility, Comprehensibility, and Accentedness

TERM	MEANING	INDEX	NON-NATIVES
ACCENTEDNESS	How strong the foreign accent sounds (0 = no accent)	Native listener	"He speaks English with a heavy Algerian accent"
INTELLIGIBILITY	How many words the listener actually understands	Any listener	Listener writes (80 %) of the words correctly
COMPREHENSIBILITY	How much effort the listener needs to understand	Any listener (rating) (1: very easy, 7: impossible)	"I understood every word, but it was difficult"

2.5.3. Assessment of Pronunciation Proficiency

Assessment of pronunciation proficiency is a crucial yet complex component of language evaluation. It involves measuring both segmental accuracy and suprasegmental control in a way

that reflects communicative competence. The development of reliable tools and the use of technology have transformed pronunciation assessment, making it more objective and more learner-centred. There three methods of assessment for learners' pronunciation as follows:

Table 2.9

Methods of assessment for learners' pronunciation

METHOD	HOW IT WORKS	FEATURES	DRAWBACKS	IN THIS STUDY
Human Raters	Trained teachers listeners give scores on a scale (1–9)	Real-life feeling includes fluency	Subjective slow expensive	Yes
Acoustic Analysis	Computer measures	(100 %) objective, exact statistics	Needs training, does not measure "feeling"	Yes (Praat software)
Automatic Speech Recognition **(Applications)**	ELSA Speak, Speechling give score 0–100	Fast, private, real-time feedback	Still not perfect for strong accents	Yes (during treatment for practice)

2.5.4. Tools and Rubrics for Measuring Pronunciation

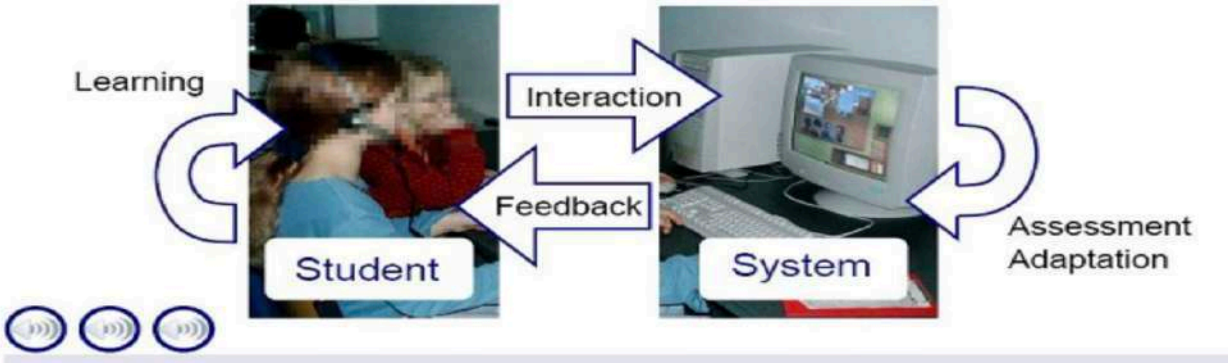
Traditional assessment tools such as oral interviews, reading passages, and word lists, remain present in evaluating learners' pronunciation skills. Rubrics typically address features like segmental production, stress, rhythm, and intonation (Isaacs, 2014). Well-designed rubrics, such as those used in the IELTS or TOEFL speaking sections, provide standardized criteria that help teachers assess pronunciation objectively. In classrooms, analytic scales focusing on intelligibility and fluency are even more effective, allowing teachers to provide targeted feedback on learners' strengths and weaknesses.

2.5.2.1. Technology-Driven Assessment Methods

Recent advances in speech technology have introduced new possibilities for pronunciation assessment. Computer-Assisted Pronunciation Training (CAPT) systems, for instance, use speech recognition and acoustic analysis to evaluate learners' pronunciation automatically (Neri et al., 2008). Tools such as Praat and automatic speech evaluation apps provide learners with visual feedback on pitch, intensity, and formant patterns, facilitating self-assessment and autonomous learning. These methods increase the reliability of pronunciation assessment while promoting learner engagement through interactive feedback. In Algerian classrooms, where teacher workload and large class sizes often hinder detailed evaluation, such technologies can make pronunciation assessment more efficient and individualized.

Figure 2.12

An Example from Computer-Assisted Pronunciation Training (Adapted from Hacker, 20099)



2.6. Teaching English Pronunciation in Algerian Context

The teaching of English pronunciation in Algeria occupies a crucial position within the broader framework of EFL education. Despite the growing importance of English as a global language, pronunciation remains one of the most challenging skills for Algerian learners. This difficulty stems from a combination of historical, sociolinguistic, and pedagogical factors, as well as from limited exposure to authentic English input. The present chapter examines the contextual and theoretical dimensions shaping pronunciation instruction in Algeria, addressing the linguistic landscape, the influence of age on pronunciation acquisition, cross-linguistic phonological variations, and the potential of multimedia tools to enhance pronunciation teaching and learning.

2.6.2.1. Historical and Sociolinguistic Factors

The teaching and learning of English in Algeria have evolved within a complex historical and sociolinguistic context shaped by colonial heritage, language policy, and globalization. Following independence in 1962, Algeria adopted Arabic as the national and official language, while French retained a strong sociolinguistic presence due to its colonial history. English was introduced later as a *foreign* language, initially occupying a secondary position in the educational

system. However, in recent decades, English has gained prominence as a language of science, technology, and international communication, reflecting a broader shift toward global engagement (Benrabah, 2014). The Algerian linguistic environment is thus *multilingual*, characterized by the coexistence of Arabic, Berber, and French, with English functioning as an additional foreign language. This multilingualism creates both opportunities and challenges for EFL learners. On one hand, exposure to multiple linguistic systems can foster metalinguistic awareness; on the other, it often results in cross-linguistic interference, particularly in pronunciation (Bouhadiba, 2018). Learners' phonological habits are influenced by their first and second languages, which shape the way they perceive and produce English sounds.

Table 2.10

Historical and Sociolinguistic Context of English in Algeria

Aspect	**Key Apsects**
Historical Background	Post-independence language policy prioritized Arabic, while French retained strong influence. English was introduced later as a foreign language.
Current Status of English	Increasingly valued for science, technology, and global communication; still limited in everyday use.
Linguistic Environment	Multilingual setting with Arabic, Berber, and French shaping learners' pronunciation and language learning.
Main Challenge	Cross-linguistic interference affecting oral proficiency and pronunciation accuracy.
Pedagogical Implication	Greater need for innovative, multimedia-based EFL instruction to enhance pronunciation and exposure.

In Sociolinguistics, English in Algeria carries symbolic value as a language of modernization and global mobility. It is increasingly viewed as essential for academic advancement and employability. Nevertheless, its use remains largely confined to academic and professional domains, with limited presence in everyday life. This restricted exposure impacts oral proficiency and pronunciation accuracy, underscoring the need for pedagogical innovation in EFL instruction.

2.6.2. Challenges in English Pronunciation Instruction

Teaching pronunciation in the Algerian context faces serious pedagogical, institutional, and linguistic challenges. At the pedagogical level, pronunciation often receives limited attention in the classroom, as teachers focus more on grammar and vocabulary to meet syllabus requirements. This marginalization results in inadequate practice opportunities for developing segmental (individual sounds) and suprasegmental (stress, rhythm, and intonation) features of English (Derra, 2020). In most Algerian institutions, large class sizes, insufficient phonetic training for teachers, and lack of technological resources hinder effective pronunciation instruction. Many teachers rely on traditional approaches, focusing on imitation and repetition rather than communicative strategies. The absence of multimedia pronunciation tools further restricts learners' exposure to authentic English input, limiting their ability to internalize native-like pronunciation patterns (Kaci, 2022).

From a linguistic point of view, the influence of Arabic and French phonological systems often leads to systematic pronunciation and stress errors. Moreover, the sociolinguistic reality of English being a language used only in the classroom means learners rarely receive natural input outside formal settings, reducing opportunities for practice and spontaneous production.

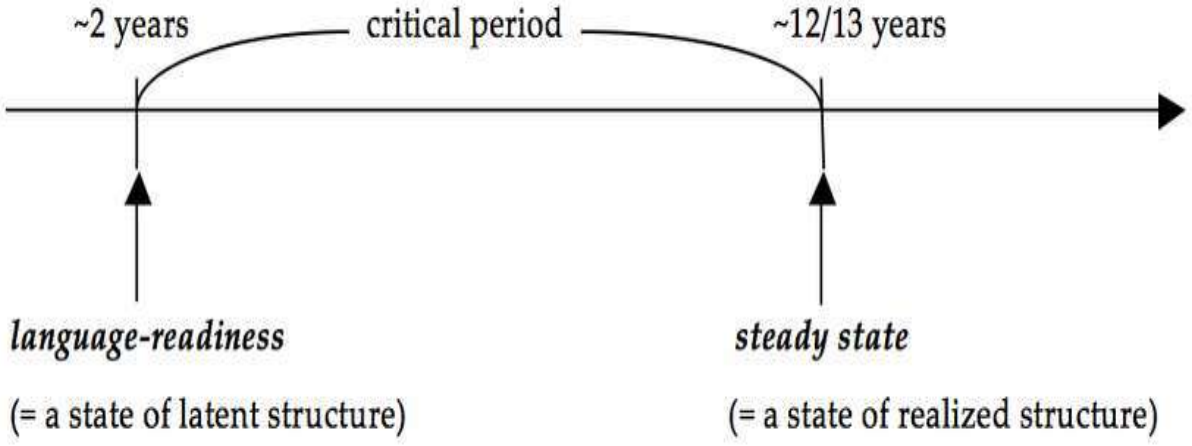
Addressing these challenges requires an integrated approach that combines explicit phonetic instruction, technological tools, and continuous teacher training. Incorporating multimedia and interactive pronunciation applications can help overcome exposure limitations and foster learner autonomy: an essential step toward achieving intelligible and confident oral production in English.

2.6.3. Critical Period Hypothesis in Pronunciation Development

The Critical Period Hypothesis (CPH) suggests that there is a biologically determined window during which language acquisition occurs most efficiently, particularly in terms of pronunciation and phonological development. Originally proposed by Lenneberg (1967), the hypothesis suggests that after puberty, the brain's neuroplasticity declines, making the attainment of native-like pronunciation significantly more difficult. This idea has been widely explored in both first and second language acquisition research and remains a key theoretical framework for understanding pronunciation challenges among adult EFL learners.

Figure 2.33

Lenneberg's Critical Period for language development



In the Algerian context, where English used to be introduced relatively late, typically during middle and secondary school, learners begin formal instruction after the proposed critical period. Consequently, they face greater difficulty in mastering the phonetic and prosodic features of English. Studies in second language phonology (Flege, 1995; Scovel, 2000, among others) indicate that early exposure to the target language facilitates the development of accurate phonetic representations, while late learners tend to retain traces of their accent from first language. This aligns with classroom observations in Algeria, where learners often exhibit persistent segmental errors and suprasegmental deviations (e.g., misplaced stress and intonation patterns).

However, more recent perspectives view the CPH as gradual rather than absolute. Scholars such as Birdsong (2018) and Muñoz (2014) argue that although early exposure provides advantages, adults can still achieve high levels of intelligibility through targeted training, motivation, and increased input. This shift has important pedagogical implications: pronunciation instruction should not aim for native-like accent imitation but rather for intelligibility and comprehensibility, goals that are attainable at any age through effective pedagogical strategies.

Incorporating multimedia tools, such as phonetic visualization software, speech recognition programs, and audiovisual, materials can help age limitations by enhancing auditory discrimination and articulatory awareness. Multimedia pronunciation instruction provides learners with repeated exposure to authentic speech, immediate feedback, and visual representations of sound patterns, all of which support ongoing phonological refinement. Ultimately, while biological constraints may shape pronunciation outcomes, pedagogical intervention and technological mediation can significantly enhance late learners' ability to produce clear, intelligible English speech. In contexts like Algeria, where early exposure is limited, integrating multimedia tools

within pronunciation pedagogy offers a practical response to the challenges posed by the Critical Period Hypothesis.

2.7. Contrastive Analysis of Phonological Systems

The accurate production of English sounds poses particular challenges for Algerian learners due to the linguistic distance between English and the languages they commonly speak, namely Arabic and French. Understanding the specific phonological contrasts and points of interference provides valuable insight into the persistent pronunciation errors observed in Algerian EFL classrooms. This section examines these challenges by focusing first on the structural differences between English and Arabic sound systems, followed by an analysis of how French phonological patterns further influence Algerian learners' English pronunciation.

Arabic and English exhibit fundamental differences in their sound systems, leading to predictable interference patterns. In the vowel system, the inventory size is little with 3–6 vowel phonemes called movements (e.g., /a/, /i/, /u/ or silent), whereas English has 20 vowel sounds (including monophthongs and diphthongs). This results in difficulties distinguishing English vowels like /æ/ vs. /ɑː/ (e.g., "cat" vs. "cart") or /ɪ/ vs. /iː/ ("bit" vs. "beat") (Bouchhioua, 2018).

Another difference, vowel reduction in English employs schwa /ə/ in unstressed syllables (e.g., "ago" /əˈgəʊ/), a feature that is absent in Arabic. Algerian learners often overemphasize unstressed vowels, disrupting English rhythm (Altaha, 2020). On the other hand, consonantal challenges in some consonants with voicing contrasts. First, Arabic lacks the /p/ vs. /b/ distinction, leading to substitutions (e.g., "park" pronounced as "bark"). Similarly, the English /v/ is often replaced with /f/ or /b/ (Benrabah, 2013). While others do not exist at all like /ŋ/ and the same for consonant clusters in Arabic syllable structure (CVC or CV) avoids complex clusters, causing

learners to insert vowels (e.g., "school" → /sɪkuːl/) or drop consonants (e.g., "texts" → /teks/) (Celce-Murcia et al., 2010).

Prosodic differences in stress and rhythm, since Arabic is syllable-timed, with equal stress duration, while English is stress-timed. Learners may produce monotonous speech or misplace word stress or they stress the consonant itself not pronouncing the syllable with prominence (e.g., "PHOtograph" vs. "phoTOgrapher") (Altaha, 2020). Concerning intonation, Arabic uses pitch variation for emphasis rather than grammatical function, leading to flat or inappropriate intonation in questions or statements (Bouchhioua, 2018).

The influence of French phonological system can create another challenge, as French, a colonial legacy and lingua franca in Algeria, it further complicates English pronunciation some phonemes are made silent like /h/ or they pronounce /g/ in the gerunds (-ing) and /s/ as /z/ in intervocalic positions.

For false friends, French and English share some phonemes (e.g., /ʃ/ in "shampoo"), but learners may overgeneralize French sounds. For example, the French uvular /ʁ/ replaces the English alveolar /r/, producing a guttural "red" (Benrabah, 2013). Additionally, the nasal vowels in French like (e.g., /ɑ̃/ in "vent") lead to hypernasalization of English vowels, as in "man" pronounced as /mɑ̃n/ (Altaha, 2020). In syllable structure and liaison, French often elides final consonants (e.g., "petit" /pəti/), which transfers to English (e.g., "cat" → /kæ/).

Or they link between words (e.g., "les amis" /lezami/) may cause intrusive /z/ or /t/ in English phrases like "law(r) and order" (Celce-Murcia et al., 2010). Likewise, French stress is fixed on the final syllable, leading to misplaced stress in English polysyllabic words (e.g., "adVERTisement" vs. "adverTISEment") and in French pitch range is narrower contributes to monotonic English speech, affecting pragmatic functions like sarcasm or surprise (Bouchhioua,

2018). In this case, consequently, Algerian learners due to their exposure to French in early age before English, their pronunciation errors often reflect hybridized interference from both Arabic and French in vowel substitutions as in /e/ (e.g., "day" as /de/) may merge with Arabic-influenced vowel shortening. Together with consonant cluster simplification in Arabic-induced epenthesis (e.g., "film" → /fɪləm/) coexists with French final-consonant deletion. Finally, rhythmic hybridity of syllable-timing from Arabic and French creates a "staccato" English rhythm, diverging from native-like stress patterns (Altaha, 2020).

2.7.1. English and Arabic Articulation Challenges

Pronunciation difficulties among Algerian EFL learners often stem from the phonological distance between English and Arabic. According to Contrastive Analysis Hypothesis (CAH) (Lado, 1957), errors in second language pronunciation frequently arise from negative transfer or interference from the native language. Arabic and English differ substantially in their **consonant and vowel inventories**, syllable structure, and stress patterns, leading to predictable patterns of mispronunciation. In terms of consonants, several English phonemes such as /p/, /v/, /ŋ/ have no equivalents in Arabic. Consequently, Arabic-speaking learners often substitute /p/ with /b/ (“pat” → *bat*), /v/ with /f/ (“very” → *ferry*), and /ŋ/ with /n/ (“sing” → *sin*). Similarly, the lack of voiced and voiceless interdental distinctions leads to substitutions in some dialects like /θ/ → /s/ or /t/, and /ð/ → /z/ or /d/ (Yule, 2016).

The vowel system of Arabic also proves to be challenging. English has a complex set of short and long vowels, as well as diphthongs, which differ from the relatively simple Arabic vowels. Learners often fail to distinguish between minimal pairs such as “ship” and “sheep,” or “full” and “fool.” Moreover, Arabic’s syllable-timed rhythm contrasts with English’s stress-timed rhythm, making it difficult for Algerian learners to produce natural prosody and stress placement

(Celce-Murcia et al., 2010). These difficulties underscore the need for explicit phonetic instruction and auditory discrimination training in Algerian EFL classrooms. Multimedia resources, such as interactive phonetic charts, pronunciation software, and visual articulatory models, can be particularly effective in addressing these segmental and suprasegmental differences through auditory and visual feedback.

2.7.2. Phonological Interference Among Algerian Learners

Beyond Arabic interference, Algerian learners are also influenced by French phonology, given the country's historical and educational bilingualism. Many Algerians acquire French as a second language before learning English, which means their English pronunciation is shaped by both Arabic and French sound systems. French phonological transfer often appears in the realization of certain consonants and vowels. For example, the French uvular /ʁ/ may influence learners' pronunciation of the English /r/, resulting in a guttural articulation rather than the alveolar approximant used in English. Similarly, the final devoicing rule in French may cause learners to pronounce voiced final consonants as voiceless (e.g., "bag" → [bak]). The nasalized vowels of French also interfere with the production of English oral vowels, leading to non-native-like resonance patterns (Tranel, 1987).

Suprasegmental aspects also reflect French influence as Algerian learners may transfer French intonation and syllable timing, producing a flatter pitch contour and equal stress distribution across syllables. This often results in speech that sounds monotonous or lacks the rhythmic dynamism typical of native English (Derwing & Munro, 2015). Addressing these interferences requires contrastive phonetic awareness, enabling learners to perceive subtle articulatory and prosodic distinctions. Integrating multimedia-based pronunciation instruction, such as audio models, visual waveform comparisons, and speech recognition apps, can support

learners in overcoming both Arabic and French transfer effects, ultimately leading to a more natural English pronunciation.

2.8. Multimedia-Based Pronunciation Instruction

Multimedia-based pronunciation instruction (MBI) means using technology to teach pronunciation in a smarter, more effective way. Instead of only teacher drills and dictates, we use videos, applications, and computers to give students real models, instant corrective feedback, and personalised practice. The table below summarises the main features of MBI as a solution to the problems and difficulties that encounter learners and teachers in dealing with pronunciation in large classes.

Table 2.14

Features of Multimedia-Based Pronunciation Instruction

ALGERIAN CLASSROOMS	TRADITIONAL FIX (SLOW)	MBI FIX (FAST + FUN)
Crowded Classes (over 50 Students)	Teacher corrects one by one	Mobile Apps & Software can give feedback to all at once
No Native Models in Class	Teacher (with accent) as only model	BBC videos YouTube
Students are Shy To Speak	Practice in front of class	Record on phone private score no reluctance
No Time To Measure Progress	Guess if students improved	Praat shows exact numbers (formants, pitch)

2.9. Technology Integration in Algerian Classrooms

In recent years, the integration of technology into Algerian classrooms has become a national educational goal aimed at improving teaching quality and learner engagement. In English as a Foreign Language (EFL) instruction, multimedia tools such as interactive software, language learning applications, and audiovisual resources have been introduced to enhance learners' pronunciation and listening comprehension. Technological integration in Algeria remains uneven, with urban schools often having better access to digital infrastructure than those in rural areas. Teachers, therefore, face challenges related to insufficient training, limited resources, and a lack of institutional support for technology-based pedagogy. When effectively implemented, multimedia enhances pronunciation learning by exposing students to authentic models of speech and providing immediate auditory feedback. For instance, software like Praat enable learners to visualize and correct their articulation patterns. Research in the area shows that audiovisual input, including movies, podcasts, and pronunciation videos, promotes phonetic awareness and helps learners internalize segmental and suprasegmental features of English. In the Algerian context, such tools can bridge the gap between theoretical knowledge and communicative competence, offering a more immersive experience that traditional instruction often lacks.

2.10. Positive/Negative Transfer Among Learners

Algerian learners of English typically operate in a multilingual environment where Arabic, French, and local dialects interact in complex ways. This linguistic diversity can both facilitate and hinder pronunciation learning, depending on how learners transfer phonological features from one language to another. Positive transfer occurs when similar sounds or articulatory habits in the first or second language (L1 or L2) help learners produce English sounds more accurately. For example, certain consonant clusters or vowel qualities shared between French and English make

pronunciation easier for students familiar with French phonology. Negative transfer is equally common. Learners often substitute English sounds with the closest equivalents from Arabic or French, leading to systematic errors such as devoicing final consonants or mispronouncing the English /p/ as /b/. As noted by Flege (1995), these cross-linguistic influences stem from learners' tendency to perceive new sounds through the filter of familiar phonological categories. In Algeria, where exposure to native English input is limited, such interferences can persist without corrective feedback. Multimedia resources; especially, those that combine visual, auditory, and articulatory cues can mitigate these issues by helping learners distinguish subtle phonetic contrasts and develop more native-like pronunciation patterns.

Conclusion

The present chapter elaborated a concise literature review on the basic concepts of pronunciation teaching and learning in the digital era. This chapter revisited the main language learning theories from behaviourism and cognitivism to constructivism and connectivism arriving at different type of feedback. Then, we moved to the evolution of pronunciation pedagogy from the traditional paradigm of native-like speech as in audiolingual approach tenets to present day communication approach focusing on recent research in intelligibility and comprehensibility, where suprasegmentals matter even more than perfect segments. In Algerian context, the EFL learners' interference in Arabic and French, late exposure and limited authentic input create persistent barriers. The next chapters will detail the methodology section of the study.

Chapter Three:

Methodology

Chapter Three: Methodology

Introduction

In the present chapter, we are going to review the basic principles and tools used in data collection and analysis in order to come up with consistent results and reliable interpretation. We utilised a mixed-methods approach in our research design to integrate both quantitative and qualitative research techniques to augment the depth and validity of the outcomes of our study. Meanwhile, to answer the research questions we set out in the beginning of the study to validate the hypotheses we formulated. The adopted design of the aforementioned methods is grounded in a practical model, permitting the researcher to select the most appropriate data analysis tools that best suits fulfilling the main objectives of the study. When it comes to analysing the EFL learners' productions in pronunciation, Praat software has been widely used, "Speech analysis programs such as Praat provide an acoustic analysis of the speech signal, often in the form of a visualization of the waveform, spectrogram, and pitch contour.

3.1. Research Design

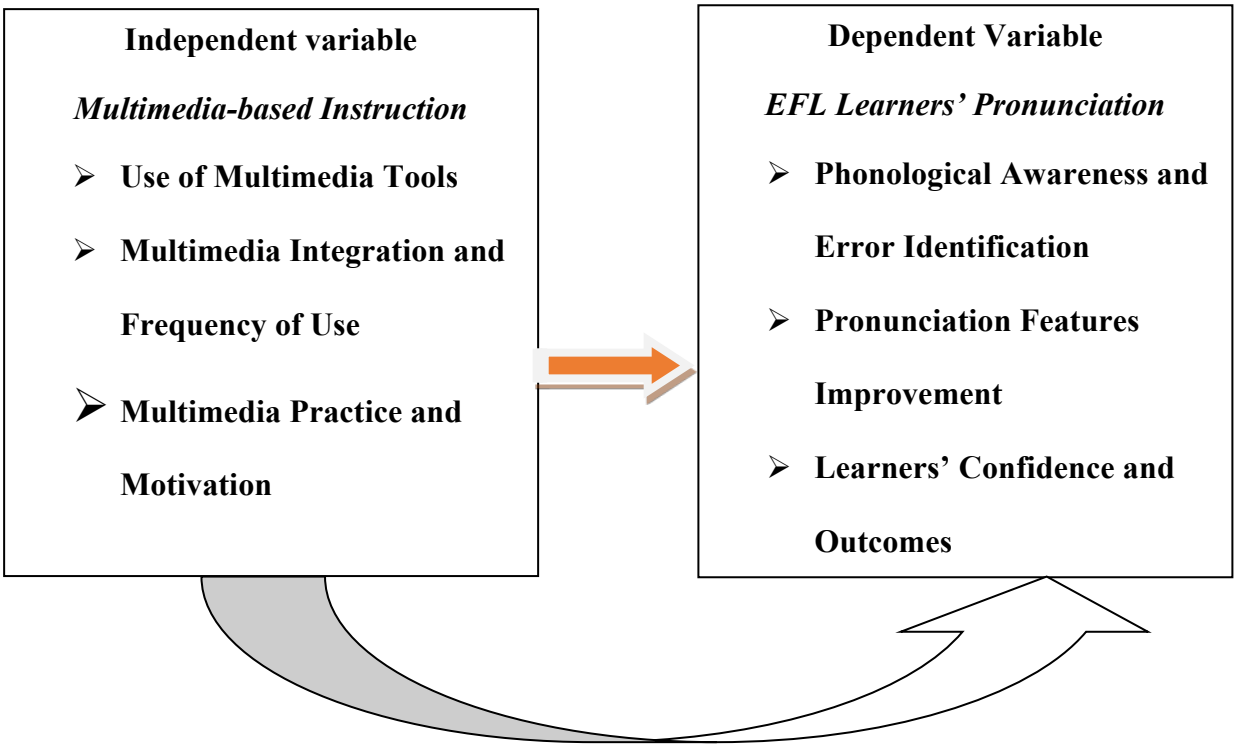
In the current study, we used a mixed-methods approach integrates both qualitative data to provide an insightful understanding of the phenomenon of mispronunciation among learners and their viewpoints and experiences towards the issue as a diagnostic stage in finding the appropriate strategies to overcome it and quantitative research methods to enhance the depth and validity of findings by analysing speech productions of the learners acoustically, responding to complex educational research questions we put forward in the outset of the study because only one method is insufficient.

3.1.1. Study Variables

The conceptual model of the study investigates the effects of multimedia-based instruction as the independent variable on the pronunciation skills of EFL learners as the dependent variable. It posits that the strategic use of multimedia tools, their frequency of integration, and the motivational practice they provide directly influence key learning outcomes. These outcomes encompass the learners' development of phonological awareness, the measurable improvement of specific pronunciation features, and a subsequent boost in their speaking confidence and overall proficiency in pronunciation and fluency. Essentially, the model frames cause-and-effect relationship where targeted multimedia intervention is the driving force for comprehensive pronunciation gains.

Figure 3.1:

The Variables of the Study



This research framework is grounded in the pragmatic paradigm, allowing researchers to select strategies best suited for the study's objectives while embracing methodological flexibility. The decision behind selecting this research design is to provide holistic views of the studied phenomenon; particularly, to gain valuable evidence in exploring language education problems in the EFL contexts. As asserted by Dörnyei (2007), in applied linguistics research methodology, he emphasized the crucial value of mixed-method approach by integrating both quantitative and qualitative data at one or more stages of the research h process and focuses on the purpose of such research, the compatibility of different research paradigms. He adds ″Quantitative surveys, when complemented by qualitative data sources such as interviews or open-ended questionnaire responses, can provide more nuanced and comprehensive insights into complex educational phenomena."(Dörnyei, 2007). The current study employs correlated schemes, with both constituents but tackled separately before being integrated for a combined analysis, ensuring comprehensive insights and methodological rigor. As illustrated below:

Table 3.1:

The Adopted Reseach Design of the Study

EMBEDDED MULTIMEDIA INTERVENTION	
NATIVE-SPEAKER MODELS	**LEARNERS' QUESTIONNAIRES**
(Pre-test → Treatment → Post-test)	**35-item students' questionnaire**
Participants' recordings (n=20)	**201 students' questionnnaire**
DATA ANALYSIS	
Praat analysis	**Questionnaire data**
Methods of Data Analysis	
ACOUSTIC DATA MEASURES	**SPSS DESCRIPTIVE STATISTICS**
DISCUSSION AND RESULTS	

To be more specific, the pragmatic approach adopted in this study seeks to answer the research questions of the present thesis in a quasi-experimental design and a survey elaborated to bring about reliable and trustworthy results of the learners' real outcomes in order to scrutinise their deficiencies and imperfections in terms of intelligibility and comprehensibility in English pronunciation with a special focus on prosodic features.

The upcoming stages report the context of the study and the background of the participants and their distribution in both the survey and the acoustic training of the control and the experimental groups. For (Schwieter and Benati, 2019), experimental research was somewhat odds in qualitative research within postpositivist paradigm by generating and testing hypotheses. Using the example of authentic listening materials on listening comprehension which means under ideal circumstances the researchers should be able to determine the effect of the intervention on learners' outcomes.

In the same regard, the adopted research design of the study involved several components listed above are arranged in a flow that starts with native-speakers' models to improve the segmental and suprasegmental features on learners' productions using both acoustic analyses using Praat and surveying learners' viewpoints and experiences through SPSS then to report them in the results and the interpretation of findings. To embrace methodological flexibility and valid conclusions, we opted for a quasi-experimental study in a longitudinal stage to track the students' development in pronunciation learning before and after the period of training to demonstrate causal relationships, in our case, the effects of multimedia-based instruction using authentic materials on the learners' pronunciation improvement to reach fluency as an optimal goal of the study. This type of analysis offers objective feedback to help students improve their pronunciation and raise their awareness of prosodic features." (El-Garawany, 2021). Subsequently, mixed-methods design

provides holistic views of the investigated linguistic and phonetic phenomena in foreign language pronunciation, particularly by examining the language outcomes and the best strategies to overcome these problems in the EFL contexts.

As it has been mentioned in the literature, that EFL learners and more specifically the Algerian learners in the course of their learning English, they face plenty of challenges that affect their mastery and fluency in the target language as mentioned in (Ghounane, 2018, p. 426) "EFL learners encounter hindrances in learning correct pronunciation due to the lack of interaction with native speakers; as a matter of fact, they need a leg up from teachers who should provide them with enough and intensive practice"

3.1.2. Study Context

The current study was conducted within Algerian university EFL classrooms, involving undergraduate learners from a representative sample out of the whole population from First- year students at the University of Biskra, in which we selected via purposeful sampling to ensure variety of viewpoints and experiences in prior knowledge of English pronunciation and its basics in telling the difference between segmentals and suprasegmentals.

The sample size was determined based on a target population of all the second-year students. So, to determine the proper sample size, we used the Robert Mason equation, which is a statistical formula designed to calculate the optimal sample for survey-based research. The aim of this procedure is to limit the total population size (N = 450, as indicated in the administration records). Then, after applying the Mason formula, which uses the following values:

$$n = \frac{M}{[S^2 \times (M-1) \div pq] + 1}$$

Where:

- M is the population size (450)
- S is the standardized value for confidence (for 95%, $S = 1.96$)
- p is the proportion expected to give the characteristic studied (e.g., 0.5)
- q is the remaining proportion (1-p, here 0.5)

This calculation recommended selecting 207 participants for your sample, ensuring representativeness and statistical validity for your findings.

The results of the equation above when calculating our whole population showed that optimal number of the recommended sample size is approximately 207 given that the full number of the population is 450 students to ensure a representative statistical validity for our findings. Therefore, after administering the questionnaires to the whole population, we collected from students 220 copies but after scrutiny, we excluded 18 copies for the reasons of inconsistency in answers then we ended up with 201 participants involved in the study.

In our case, the probability sampling was chosen carefully to obtain reliable results in which the researcher's judgment not to select non-probability sampling but rather every informant of the population was chosen on purpose to avoid any kind of chance in the probability sampling to minimise bias and to get generalisable results.

The sampling process of selecting the convenient subset of individuals or units in the study from a larger population so that conclusions can be drawn about the whole, without studying every member. The adopted method was based on Singh's typology, as a sampling method to support representativeness and valid statistical inference in real-world settings as follows:

Table 3.2

Types and Methods of Sampling (Adapted from Singh, 2006)

Method	Description	Use Case Example
Simple Random	Equal probability for all units (e.g., random number generators).	National health surveys
Systematic	Select every *k*-th unit from a sampling frame (e.g., every 10th student).	Quality control in manufacturing
Stratified	Divide population into subgroups (strata).	Comparing gender-based income disparities
Cluster	Randomly select clusters (e.g., schools).	Educational research in rural areas
Multistage	Combine methods (e.g., cluster → stratified → random).	Large-scale demographic studies

Note. Adapted from (Singh, 2006)

In the second stage, we used a speech analysis software called Praat(TM) to analyse the productions of the students in articulating sounds and words' pronunciation for accuracy to improve EFL learners' intelligibility and comprehensibility. Henceforth, the EFL students who participated in the training sessions, which were divided into two groups. The first group was assigned as a control group, which has learnt using the traditional method of teaching in English pronunciation for eight weeks in a duration of twelve hours of intensive instruction, whereas the other group which is experimental group that has been exposed to English pronunciation using the

same content with embedding multimedia resources in the classroom during the same period of instruction (MBI teaching).

Figure 3.2

Functions of Praat Software

Items	Functions
Waveform	It is the top half of the Sound Editor window. You can visualize the wave and draw it to the Praat Picture window.
Spectrogram	It is the bottom half of the Sound Editor window. It displays the acoustic characteristics of speech- the formants, pitch contour, duration and intensity and provide detailed information about them.
Voice bar	It is the dark bar in the spectrogram. The darkness of the voice bar shows the intensity of the sound. This feature can be used to separate /p, t, k/ from /b, d, g/.
Pitch	In the Spectrogram, the blue line stands for the pitch's rising and falling. It can be used to see the stress, tone of the word and intonation of the sentence.
Formant	In the Spectrogram, the red dots represent formants. When analyzing the vowels, we would focus on the first three formants: F (1), F (2) and F (3). It can reflect the place and manner of articulation of vowels.
Intensity	The intensity is marked by darkness of the bands in the waveform, and marked as a yellow line in the spectrogram.
Duration	Duration is quite an essential property of the speech sounds. It can be used to compare the time that a specific sound or pause costs, e.g. /p, t, k/ show longer closure duration than /b,d,g/.

Note. Adapted from (Boersma & Weenink, 2025)

The overall all number of the subjects of the study that have had pre-tests and recordings while pronouncing English on phoneme-level and prosodic-level proficiency were forty

participants divided in two groups of twenty subjects as indicated earlier before the experimentation and after the treatment as a post-test during the different phases of the experiment. Then, the final scores of the students were compiled to be compared with the native-speakers' audio recordings using the acoustic analysis software Praat(TM). The latter is commonly used in EFL pronunciation research and monitoring accuracy to facilitate speech analysis by providing visual feedback on acoustic descriptions such as pitch, formants, duration, and waveform.

In teaching pronunciation, Praat can help EFL teachers and learners visualize pitch variations and intonation patterns via speech waveforms and pitch contours, which significantly improves their ability to produce correct utterances and intonation in English speech. A quasi-experimental study was conducted by (Bengrait, 2020, p. 32) showed that students using Praat had significant improvement in EFL intonation production compared to control groups receiving traditional instruction. On the international scope, (Gorjian et al, 2013) proved that using Praat software in teaching prosodic features to EFL Learners when practicing stress and intonation through CALL approach were more successful than the students who were taught through traditional method. In the same vein, a recently published article by (Nguyen, et al, 2025, p. 84) revealed significantly higher post-test pronunciation scores in the experimental group compared to the control group. Additionally, the "students perceived SpeechAce as a user-friendly and effective tool, particularly praising its immediate feedback and interactive features". Accordingly, learners and teachers alike can use Praat to visualize and compare speech productions in contrast with native speaker models, enhancing their awareness and control over pronunciation aspects.

3.1.3. Participants' Profiles

Before starting collecting data, to avoid any kind of threats to the study's validity and reduce research bias and respondents' subjectivity, we opted for both a questionnaire and a quasi-experimental study in mixed-methods design to ensure the feasibility and reliability of the research outcomes. The study involved two distinct participant groups whose demographic and linguistic profiles in the consent forms (see appendix) are summarised below:

Table: 3.3

Participants' Demographic and Linguistic Profiles

Gender
Age (Years)
Secondary-School Stream
Years Studying English
Self-Rated Overall English Proficiency (1–5)
Self-Rated Pronunciation Level (1–5)

A total number of 20 first-year LMD students (10 Control Group, 10 Experimental Group) were recruited from two intact classes of the module "Linguistics and Phonetics" at the Department of English, Mohamed Khider University of Biskra, throughout the second semester of the academic year 2024–2025. The high proportion of female respondents reflects the gender distribution typical of English departments in Algerian universities (approximately 75–80% females of the sample). This imbalance has been openly admitted as a limitation in the next chapter of data analysis but it does not invalidate the findings, given the consistent patterns across gender in the process of teaching. Both samples were recruited voluntarily, provided informed consent

(Appendix 1), and were assured full anonymity. Participation posed no academic penalty or reward, ensuring natural responses.

In this regard, questionnaires are most commonly used for collecting both qualitative and quantitative data by exploring informants' background information and closed-ended questions as qualitative instruments and quantitative data using Likert scales, and numerical data for statistical analysis in coding large scale responses. For example, the questionnaire in our case seeks to investigate the EFL learners' pronunciation on a scale from 1 to 5-points (35-items questionnaire, see appendix 3) that can be calculated with descriptive or inferential statistics. Likewise, quantitative surveys are particularly suitable for measuring tendencies across large samples allowing for reliable generalizations. (Wium & Louw, 2018).

3.1.4. Participants' Consent Form

The participants who took part in the study were requested to sign the consent form (see appendix 1) to participate in both stages of the research. First, to complete a pre-intervention survey prepared to review the personal information of the subjects in the experiment to track their progress in the pre-test, treatment and the post-test. Second, the main purpose of this survey is to design questionnaire items that reviews learners' preferences, attitudes toward multimedia-based pronunciation training, and prior exposure to digital learning tools and other dimensions. The survey was piloted to two groups of thirty students each, then it has been concluded with some rectifications and adjustments to refine item clarity and reliability. Accordingly, the final version of the questionnaire was subjected to experts in the field and designed in a bilingual structure to ensure full understanding of the items.

3.2. Choice of Method

The present study adopted a mixed-methods design in a deliberate theoretical and methodological ends due to the multifaceted, multidimensional nature of pronunciation in the Algerian context where many internal and external factors affect the results of the study. As pronunciation and multimedia influence cannot be measured employing a single method such surveying the students' attitudes and preferences or merely an experiment with a small number of participants. Therefore, to reduce any kind of bias or subjective in research and to provide firm evidence in the findings of the study we combined both methods in a mixed-methods design.

The latter method, in our study, can measure multilevel variables and dimensions such as measurable acoustic change and perceptual judgement of intelligibility and comprehensibility by raters as well as learners' individual experience of progress, motivation, and reduced anxiety. Methodologically speaking, no single method can adequately fulfil this multifaceted phenomenon. On the one hand, some research in second language pronunciation pedagogy employed only experimental studies that are only quantitative acoustic studies (Saito & Plonsky, 2019; Lee et al., 2020) but these studies did not manage to overgeneralise and to find why ESL learners still have accent and problems in pronunciation. On the other hand, purely qualitative approaches (Sardegna & McGregor, 2023) thoroughly reviewed learners' beliefs and attitudes in their studies but they cannot establish causality or measure phonetic change and gain. Thus, the choice of a mixed-methods design is adequately appropriate for the current research study.

3.2.1 Quasi-Experimental Design

Due to some institutional constraints in our context, the Algerian universities, such as intact classes, fixed timetables, and the impossibility of true random assignment across subjects, we

opted for a quasi-experimental with a small number of participants in pre-test and post-test for control-group and experimental-group design. This design preserves internal validity while maintaining natural validity in real educational settings, a combination widely accepted in applied linguistics intervention research (Plonsky & Oswald, 2017). The application this design was conducted in two parallel first-year phonetics classes (n=10 each) were designated as Control Group (CG) and Experimental Group (EG). Pre-treatment equivalence was verified through independent-samples t-tests on age, years of prior English study, and self-rated proficiency. (all $p > .05$; see Section 3.4), satisfying the key assumption of quasi-experimentation.

The learners involved in this study underwent a form-focused pronunciation teaching, with a long-term training using multimedia-based pronunciation instruction, utilising authentic audiovisual aids and interactive content to model segmental features, mouth movements, and native-like English productions in terms of intelligibility and comprehensibility to reduce accentedness. To reach a comprehensive insight into the carried research, an in-depth quasi-experimental research design was designed. Thus, the study comprises a speech analysis of the participants' productions of the eight weeks of treatment for both of the control group and the experimental group to measure if there is a significant difference between both groups or not. As suggested by Saito (2021, p. 808), after reviewing more than fifty empirical studies on L2 pronunciation teaching, he put three fundamental questions that future research can investigate in order to determine if explicit pronunciation instruction actually works in the longitudinal studies. In this study, we will answer the questions he raised concerning whether the instruction is indeed effective in the terms of **focus of instruction of** segmentals or suprasegmentals and the **type of instruction**, that is, the focus on form or on forms as well as the **type of outcome measures** either controlled or spontaneous production.

Table 3.4

The Study's Quasi-Experiment Design

Phase	***Stage***	***Focus Area***	***Activities***	***Assessment***
1	Pre-Test	Standard Pronunciation Assessment	Pre-test: Recording participants reading a set of diagnostic words and sentences covering consonants, stress patterns, and intonation (declarative and yes/no questions).	Using Praat: Analysis of the recordings for accuracy. Learners' Outcomes rated in scale per number of errors.
2	Treatment (8 weeks) Eight sessions Using audio tracks and authentic videos and	Pronunciation Instruction: Segmental sounds & Minimal Pairs. Suprasegmentals	Listen and repeat words (English phonemes). Practice minimal pairs: Read aloud after shadowing and drilling.	Daily recordings of word lists and minimal pairs; Assessment and feedback against native-speakers' productions

	PowerPoint slideshows And English Pronunciation Programme TeacherApp, ELSA Speak	Instruction method	Control Group (traditional teaching) Experimental Group (Multimedia-Based Instruction teaching)	Find the statistical differences
3	Post-Test	Final Pronunciation Assessment	Re-administer pre-test materials: Record the same diagnostic words, sentences, and drills. Compare pre- and post-recordings for improvements in segmentals and suprasegmental. Include extended reading from all weeks.	Final voice recordings; quantitative scoring using Praat results and rater's feedback on learners' progress.

3.2.1.1 Description of the Stages of the Intervention

Before the intervention, all participants signed the consent form stating the detailed report of the description of the study and the procedures of training and assessment. Then, the treatment of the learners' pronunciation employed a three-stage design. First, a pre-test for the forty subjects then a training for eight-week instruction in a traditional method for the control group and the other

group was subjected to MBI framework and finally a post-test to track progress and to find significant statistical differences. It must be noted that the obtained data from learners' pronunciation results were analysed using the speech synthesis software Praat to quantify variations in intelligibility, accuracy, and segmentals as well as the suprasegmental features. Consistently, the current design intends to investigate the effect of traditional and multimedia-based instruction on the EFL learners' pronunciation performance in terms of the aforementioned measures. The participants were 20 intermediate-level college students studying first- year at the University of Biskra in the department of English who were assigned to two groups. The students' personal details are represented in the table below:

Table 11:

Personal Details of the Participants of the Intervention

Participants	*Gender*	*Age*	*English Education (years)*
Student 1	F	18	7
Student 2	F	19	8
Student 3	F	18	7
Student 4	F	19	8
Student 5	F	20	9
Student 6	F	18	7
Student 7	F	18	7
Student 8	F	17	7
Student 9	F	18	7
Student 10	M	19	8
Student 11	F	20	9

Student 12	F	18	7
Student 13	F	18	7
Student 14	F	19	8
Student 15	F	18	7
Student 16	F	19	8
Student 17	F	18	7
Student 18	F	18	7
Student 19	F	18	7
Student 20	M	19	8
Average		18.45	7.5

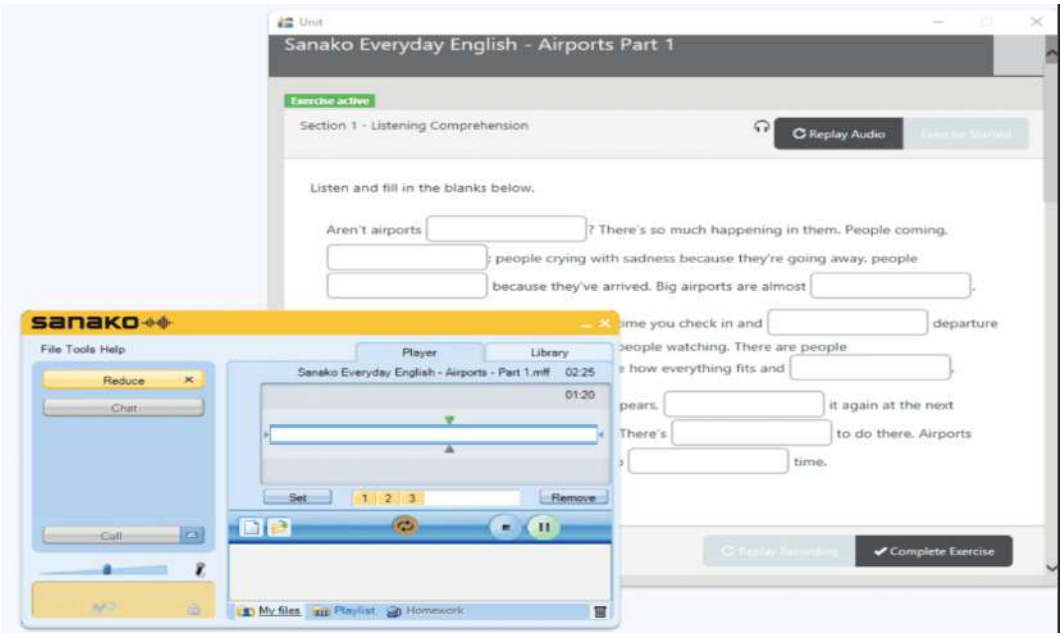
Under an eight-week course, the participants learnt the basic of pronunciation using either embedded multimedia authentic materials or the traditional way of teaching without integrating digital content. Adopting a quasi-experiment design, to measure within-group members differences and between both groups comparisons were made statistically. Consequently, paired samples t-tests were calculated to compare the pretest and post-test of each group (within-group assessment) to measure any improvements between the pretest and the post-test scores. Likewise, we analysed the results among the post-test of the two groups (between CG and EG differences). The outcomes of these tests in either traditional or digital mode will show the effects of embedded multimedia on students' overall fluency. Moreover, the final results will be consolidated with the analysis of the data obtained from the post-intervention questionnaire.

3.2.1.1.2. Analysis of the Learners' Productions using Praat

The learners' records were collected from a representative set of 20 (males = 2; females = 18) adult EFL students majoring in English, aged between 19 to 25 years-old. The subjects of the study were volunteering students from two groups studying English, who were enrolled in a pronunciation training course at Biskra university. All the participants had a minimum of 8 years of exposure to English in their middle and high schools, and they estimated their level in the pre-intervention survey as intermediate to upper-intermediate English proficiency and the choice of their preferred accent ranged from British to American varieties whereas just a minority opted for the option of intelligibility and comprehensibility. More importantly, the treatment sessions were conducted the foreign languages laboratory equipped with SANAKO(TM) desktop and pronunciation teaching equipments as illustrated below:

Figure 3.3

A Screenshot of the SANAKO(TM) Desktop

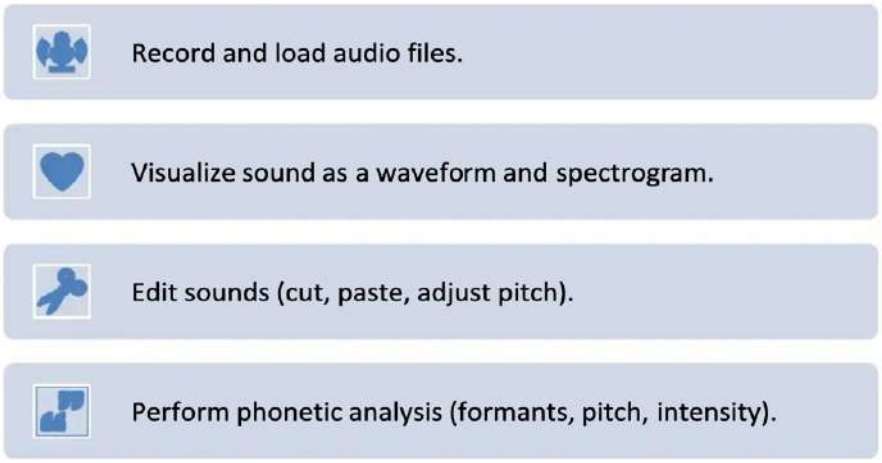


Practically, an oral test was conducted at the end of each 90-minute session where subjects were asked record their oral productions. Therefore, we, first, recorded the participants' producing the target phonemes in the selected words and sentences (single wordlists, minimal pairs or sentences containing the investigated patterns) using Boya BY-M1 microphone to get a good-quality audio files.

The analysis of the participants' oral productions show before and after-instruction stages on a par with acoustic measures for segmental (formants and vowels accuracy) and suprasegmental (pitch and stress placement) features. To do so, we used Praat software to compare accurately what students pronounce to analyse them in contrast with native-speakers' models as control words and sentences for baseline comparisons.

Figure 3.4

Basic Feature of Praat Software

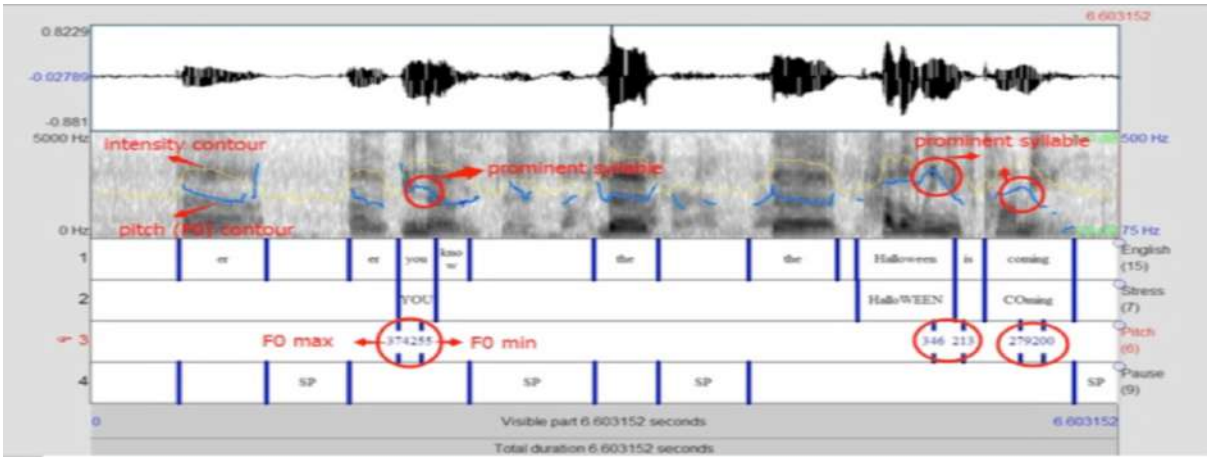


The speech analysis software Praat is used to visualise the waveforms and spectrograms of the audio samples of the participants. It is used to illustrate the acoustic analyses of spectrogram with Formant Tracks showing a post-instruction in comparison with native-speakers' norms. In

addition, the pitch contours of question–statement patterns in intonation could be measured in post-instruction to show native-like intonational outcomes. Moreover, the representation of waveforms with syllable and stress markers divided into syllables, with Praat's annotations indicating primary stress before and after treatment for viewing accuracy.

Figure 3.5

A Sample of Speech Analysis Using Praat



The main purpose of using Praat is to measure the overall validity of the content of the test and to review the construct validity by some experts in phonetics and pronunciation in order to confirm the main dimensions of the test, that is, the segmental and suprasegmental measures in comparison with multimedia contents. In other words, the reliability of the test for internal consistency could be calculated to find (Cronbach's α) through testing and retesting the subjects to assess statistical differences between the two groups using Praat (for speech analysis) in the experimental design to extract quantitative features and measures. Subsequently, the obtained scores of learners' productions in the pretest and post-test stages are analysed using SPSS following the model below:

Table 3.6

Researchers' Speech Analysis Features using Praat

Pronunciation features	*Metrics*	*Statistical significance*	*Comparison of scores (MBI vs Control group)*
Segmental improvements			
Vowel and consonants accuracy	F1/F2 with native norms Voice Onset Time (VOT) measurements	p-value t-tests	Pre- and Post-Test Mean Scores Post-test VOT accuracy
Suprasegmental improvements			
Syllable, stress, intonation	Comprehensibility ratings	Errors and their effect size	Native-like contours and means
Overall comparison	Comprehensibility Intelligibility Accentendness	Overall assessment	Estimation of the MBI participants level vs control group

3.2.2. Students' Questionnaires

Student questionnaires, in this study, were elaborated to ensure a more thorough investigation. As Dörnyei (2007, 2010) suggests, this method enriches research in three key ways: it captures student perspectives that can reveal new avenues of inquiry, it creates a more complete picture of the classroom by incorporating the learners' voices, and it generates quantitative data that can complement qualitative findings through descriptive statistics.

The use of student questionnaires constituted a central component of the present mixed-methods design, serving both exploratory and confirmatory purposes. Questionnaires were specifically developed to capture Algerian EFL learners' perceptions, attitudes, motivational orientation, and self-reported improvements following exposure to multimedia-based pronunciation instruction. As Dörnyei (2010, p.16) argues, well-designed questionnaires in applied linguistics research fulfil three main functions:

1. Discovery function: they uncover issues and constructs that researchers may not have anticipated from the literature.

2. Holistic function: provide the learners' perspective, thereby creating a more valid and of classroom processes.

3. Complementary function: generate reliable quantitative data that can be triangulated with experimental results, perceptual judgements, and acoustic measures, enhancing the robustness of mixed-methods inference (Creswell & Plano Clark, 2018; Dörnyei & Taguchi, 2010).

In pronunciation-specific research, questionnaires have proven particularly valuable for documenting affective and motivational variables that objective acoustic measures cannot capture (Baker, 2014; Nagle, 2021; Sardegna & McGregor, 2023). Learners' self-perceived improvement,

confidence improvements, willingness to communicate, and preference for certain instructional modes are now widely recognised as mediating factors in the ultimate success of pronunciation interventions (Huensch, 2019; Lee et al., 2020).

3.2.2.1 Description of the Students' Questionnaire

The current structured research questionnaire was elaborated to investigate students' perceptions and experiences regarding the use of multimedia tools for learning English pronunciation. It is meticulously organized to gather both demographic data and detailed attitudinal information through a 5-point Likert scale. At its outset, the "*Background Information*" section serves as a foundational component of the research instruments. It is designed with the purpose of collecting crucial contextualised data about the participants. By employing closed-ended, multiple-choice questions, this section ensures the efficient collection of standardized, easily quantifiable data that is crucial for subsequent subgroup in data analysis.

The specific questions targeted both gender and age **(Q1 & Q2**) which are standard demographic markers that allow the researcher to examine whether perceptions and experiences with multimedia-based pronunciation instruction are uniform across the population or if they vary significantly based on these indicators. Furthermore, the inquiry into the EFL students' secondary school stream **(Q3)** provides valuable insight into their academic preparation and orientation before enrolling in studying English. This is particularly salient in an EFL context, as students from a ***"Foreign Languages"*** stream are likely to have a different foundational exposure to English phonology compared to their peers from literary or scientific streams, potentially influencing their speaking skills and their receptiveness to the instructional method being studied.

The second section is deliberately designed to establish key standard metrics against which the main attitudinal data can be interpreted. The questions ***four*** and ***five*** probe the students' self-assessed overall English level and pronunciation skills. They create a pre-intervention profile of perceived proficiency enabling the researchers to perform a more nuanced analysis. For instance, the data can reveal whether students who self-identify as having a "***Poor***" initial level perceive a greater benefit from multimedia tools. This is complemented by (***Q6)*** and (***Q7)***, which explore behavioural and motivational precursors. The accent preference question ***(Q6)*** delves into the learners' goals, revealing whether their engagement is driven by a desire for a specific native-like model (British or American) or simply for overall intelligibility, which can influence their choice of materials and satisfaction with outcomes. Afterwards, the question on the current frequency of audiovisual aid use **(Q7)** effectively segments the cohort, distinguishing between novices and experienced users of multimedia for pronunciation training which is essential for attributing any reported outcomes directly to the study's intervention.

The questionnaire uses is meant to measure the intensity of students' attitudes across four thematic dimensions. The first dimension, **"Multimedia-based instruction"** investigates the fundamental appeal and perceived usefulness of the tools. It examines whether students prefer them over traditional methods and probes the value of specific features like the ability to visualize mouth movements and tongue positions, which are critical for accurate sound production. This dimension also explores how authentic materials and interactive software contribute to raising phonological awareness and enabling self-correction.

Building on this, the second dimension which is **"The Integration of MBI"** shifts focus from perception to practice. It examines how seamlessly these tools are incorporated into the learners' routines, assessing the frequency of use, the perceived ease of learning, and the level of motivation

they inspire. This section is vital for understanding whether the tools are engaging enough to promote consistent practice and, most importantly, to encourage autonomous learning outside the formal classroom setting.

The final two dimensions tend to measure the ultimate outcomes of this process. The third dimension which is designated as **"Pronunciation Features"** that relates the digital tools usage to improvements in specific speaking skills. It discreetly differentiates between segmental features (individual sounds) and the more complex suprasegmental features of stress, intonation, and rhythm, which are essential for natural and intelligible speech. The questions here target the practical goal of pronunciation: being clearly understood by others and reducing the barriers of a strong foreign accent. Finally, the fourth dimension, **"Learners' Outcomes"** measures the growth in learners' confidence, the reduction of anxiety, the enjoyment of practice, and the positive feedback from teachers and peers.

Table 11

The Questionnaire's Dimensions and Indicators

Main Dimensions	Indicators	Item Number
Multimedia-Based Instruction	(Use of Multimedia Tools)	1, 2, 3, 4, 5, 6, 7, 8
	(Multimedia Integration and Frequency of Use)	9, 10, 11, 12, 14
	(Multimedia Practice and Motivation)	13, 15

Pronunciation Features	(Phonological Awareness and Error Identification)	5, 6, 18, 21
	(Improving Pronunciation Features)	16, 17, 20
	(Speech Clarity and Comprehensibility)	19, 22, 28
LEARNERS' OUTCOMES	(Confidence)	23, 24
	(Outcomes and Motivation)	25, 26, 27

This section presents the statistical analysis plan for this research study. It details the methods to be used in order to test the different hypotheses. We will determine statistically the overall influence of the main independent variable (e.g., *Multimedia-based Instruction*) on the main dependent variable (e.g., *Learners' Pronunciation*). To provide a more comprehensive understanding, multiple regression will then analyse how the specific sub-dimensions of the independent variable affect the indicators of the dependent variable. Furthermore, Independent Samples *frequencies* will be employed to compare perceptions between the indicators' groups (such as males and females), while the inferential measures will be used to examine if there is significant difference in the dependent variable across multiple demographic or background categories (such as age groups or academic streams).

3.2.2.2. Piloting and Validation of the Questionnaire

The pilot study for this research was conducted as an essential preliminary step to refine the data collection instrument and ensure the validity and reliability of the subsequent main study. It was administered at Mohamed Khider University of Biskra with a small, representative sample of EFL students of two groups (60 First- year students). The initial questionnaire design deliberately used a structured format to facilitate quantitative analysis. Consequently, open-ended questions were excluded in questionnaire design, and the items were simplified to ensure full clarity for the participants. The primary goal of the pilot was to test the prepared format, evaluating the precision of the Likert-scale items, the logical flow between the sections, and the overall practicality of the questionnaire's administration.

The procedure for the pilot study was methodically carried out in two phases. First, participants were required to filling the entire questionnaire under conditions designed to trigger the main study. After this, a careful scrutiny of the collected copies was logically processed to improve the quality of the layout and the content according to the feedback on the closed-ended structures. The analysis of the returned questionnaires and participant comments revealed that certain items, including open-ended and Yes/No questions were problematic for statistical analysis and were consistently declined, indicating issues with their clarity or relevance. Based on this feedback, an analytical decision was made to eliminate these flaw question formats to rationalise the questionnaire structure for an accurate analysis of results.

The insights gained from the pilot study led to significant refinements. It results in a final questionnaire composed exclusively of multiple-choice background questions and a unified sequence of 5-point Likert-scale statements. This homogenization of the question format ensured

the data would be consistent for elevated statistical analysis. The data from the revised Likert-scale responses were analysed for internal consistency using statistical measures like *Cronbach's Alpha,* which verified the reliability of the constructs being measured. Ultimately, this rigorous piloting process ensured that the final questionnaire was methodologically coherent, clear for participants, and optimized for the quantitative research objectives.

3.2.2.3. Analysis of the Questionnaire using SPSS

The current questionnaire was analysed using ***SPSS (version 26)*** began with careful data entry, where each respondent's answers were coded into the data editor, with each column representing a survey question and for each row a participant. Clear variable names and value labels were assigned to ensure accuracy and interpretability of the results. Initial data checks were performed to identify and to address any missing values, outliers, or inconsistencies, followed by the use of descriptive statistics, (such as frequencies, percentages, means, and standard deviations), to summarize responses and explore the overall trends in the dataset.

Following data cleaning excluding any anomalies and descriptive analyses, reliability was measured using Cronbach's Alpha (coefficient α) to evaluate the internal consistency of grouped items within the questionnaire. For inferential analysis, appropriate statistical tests were conducted to examine relationships, differences among variables, to answer the research questions. The results of these analyses were then organized into tables and graphs generated by SPSS. It allows for clear interpretation and presentation of the findings in the study.

Table. 3.8

Codification of the Questionnaire Items

1	Gender	Age	X3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14
2	1	1	2	2	3	2	3	4	4	3	3	4	3	5
3	2	1	1	3	3	1	2	4	4	3	3	4	3	4
4	2	1	2	3	3	2	3	4	4	3	3	4	3	5
5	1	1	1	2	4	1	4	4	4	3	3	4	3	5
6	2	1	1	3	2	1	4	1	3	4	4	3	3	4
7	2	1	3	2	2	1	3	5	4	5	4	5	5	4
8	2	1	2	3	2	2	2	4	4	5	3	4	3	4
9	2	1	1	2	2	2	1	4	4	5	6	4	2	4
10	2	2	1	2	2	1	4	1	2	1	4	1	3	4
12	2	2	1	3	4	1	3	2	3	5	5	5	5	4
13	2	2	2	3	4	1	3	1	3	4	4	3	5	1
14	2	1	3	1	4	1	4	4	5	4	3	5	5	5
15	2	1	2	3	4	1	4	3	4	5	5	5	4	4
16	3	4	5	4	3	1	1	5	1	2	3	1	2	3
17	2	2	3	3	4	2	3	5	5	4	4	5	4	4
18	2	2	3	2	4	1	3	3	5	5	5	5	4	5

The above table is an example of how students' responses to a questionnaire have been coded numerically for entry into SPSS. Each row corresponds to an individual student's set of answers, while each column represents either a background variable (like *Gender* and *Age*) or a specific questionnaire item (labelled as X3, x4, x5, etc.). Responses that were originally categorical (for example, Likert scale options such as "*Strongly Agree*" to "*Strongly Disagree*") have been converted to numerical codes, typically ranging from 1 to 5. This coding process, known as codification, standardizes answers so the data can be effectively processed, analysed, and interpreted by statistical software such as SPSS, Statistica or R. It also ensures consistency and accuracy during analysis, enabling calculations of frequencies, means, and correlations among the different variables and student groups.

Table. 3.9

The Labels of the Used Five-point Likert Scale

Strongly disagree	disagree	Neutral	Agree	Strongly agree
1	2	3	4	5

The above table reveals the response to each statement ranges from score (1) to score (5) as a response format in the designed questionnaire about the learners' language-attitudes. Each item (statement) is followed by five ordered response options and the obtained results will be interpreted according to the guideline of the range of responses as follows (1.00–1.79) for strongly disagree label as well as the interval (1.80–2.59) for the option of disagree and (2.60–3.39) for Neutral in the questionnaire. In addition, the value (3.40–4.19) the statements replied with Agree and the range of (4.20–5.00) for the strongly agree option. Accordingly, the basic descriptive statistics we intend to report for each item in the complete questionnaire:

- Mean (M) → average score
- Median
- Mode (as the most frequent responses)
- Standard deviation (SD) how much opinions diverge.
- Percentage of respondents in selecting each item.

3.3. Ethical Considerations

Prior to the beginning of the study, formal ethical approval was secured from the relevant participants in a consent form (see appendix 1). Following this approval, the principle of informed consent was rigorously upheld. More than that, before any data collection, all participants were

provided with a detailed information sheet to inform them about the study's purpose, procedures, potential and benefits, and the anonymity of their personal data. Written informed consent was then obtained from each individual to confirm their voluntary agreement to partake in the research based on a full understanding of what their involvement would entail.

Throughout the study, the researchers implemented strict measures to protect participant rights and data integrity. Participation was entirely voluntary, confidential, and individuals were explicitly informed of their right to withdraw from the study at any point without consequence. To ensure confidentiality, all the collected data, including questionnaires and any identifying information, and recordings were anonymized during the analysis and reporting phases. Personal identifiers were replaced with codes to save participants' privacy and maintain the highest ethical standards from data collection through to publication of the results. Correspondingly, the study confronted many constraints in collecting data in the language laboratory and the voluntary participation of the students in addition to the process of anonymising recordings and data cleaning. Yet, we managed to reduce any significant random errors or bias in control and experimental groups, during piloting of the survey and the training of the participants in the experiment to impose rigour on the internal and external factors that may affect the results of the study.

Conclusion

The current chapter reviewed the methodological framework of the study and its participants from theory into practice to bring about robust and reliable data and conclusions. By explaining why employing a quasi-experimental study together with a post-intervention questionnaire to survey the students' attitudes and experiences towards the implementation of multimedia-based contents. As indicated in the chapter, the collected data was piloted, treated and the audio recording were cleaned, and completely scrutinised in every aspect to yield valid and reliable results in statistical measures. The questionnaire and Praat analyses, in turn, strengthened the trustworthiness of the study and reduced any missing values and errors to avoid any occasional contradiction between students' self-reported outcomes and the measured comprehensibility and intelligibility scores in order to gain precise findings. Ethically and practically, every decision in the choice of the methods of sampling, data collection and analysis of results was deliberately explained and justified. Ultimately, the methodology we designed was flawless and reliable to get answers to the research questions we formulated and the hypothesis we intend to confirm whether the results of the next chapter show any statistically significant differences when using multimedia-based instruction to get better scores in pronunciation.

Chapter Four:

Results and

Discussion

Chapter Four: Results and Interpretation

Introduction

This chapter presents the results obtained from analyses conducted using ***Praat software***, a specialized phonetics tool that enables the objective acoustic analysis of pronunciation features. These analyses measured formant frequencies (F1, F2) for vowel quality, voice onset time (VOT) for phoneme-level distinctions, pitch contours for intonation, and duration for rhythm and stress, providing empirical, quantifiable evidence of phonological development to raise learners' awareness in terms of intelligibility and comprehensibility to reduce accentedness. Furthermore, the research employs a comprehensive quantitative methodology utilizing a range of statistical tools analysed using ***SPSS version 26***. This statistical framework includes ***descriptive statistics*** (frequency distributions, means, standard deviations), ***reliability analysis***, non-parametric correlations (Spearman's Rho: *r*), and multiple linear regression to analyze the relationships between instructional variables and learning outcomes.

4.1. Data Analysis

The analysis of the data obtained in the current study from both instruments were revealed separately in two sections in the subsequent subheadings then the combined scores will be shown combined. The results from read-aloud pre-tests and post-tests for the experimental and the control groups and the questionnaire were analysed using Praat and SPSS respectively to create descriptive statistics to track students' learning progress before and after the intervention as well as their experiences and expectations. The overall analytical strategy employed. quantitative data (acoustic measures, perceptual ratings, questionnaire responses) and qualitative insights from open comments were collected simultaneously, analysed separately, and then merged during interpretation using joint-display tables. To get more inferences beyond a single method we added the perceptual ratings to support the statistical scores that were calculated to measure distance from

native-speaker reference values (BBC Learning English corpus). Three experienced Algerian EFL lecturers (minimum 8 years teaching pronunciation) rated all 40 recordings (20 pre + 20 post) in randomised order using 9-point scales for comprehensibility and accentedness (as mentioned in Derwing & Munro, 2015). Raters underwent a training session with benchmark samples. Intra-class correlation coefficients ICC = .87 (comprehensibility) and .85 (accentedness) which signifies an excellent reliability. For the questionnaire analysis, it was designed in a 5-point Likert scale and the overall focus of data analysis is to calculate the following:

- Descriptive statistics (means, SD, frequencies)
- Cronbach's α for reliability
- Spearman's rank correlation (ρ) – because data were ordinal
- Multiple linear regression (to see which factors best predict pronunciation improvement)
- Independent-samples t-tests / Mann-Whitney U for gender and other group comparisons

4.1.1 Analysis of the Experiment Results

This section of data analysis introduces and details the process of interpretation and discussion of created from the software's output to find the segmental and prosodic features of the participants' pronunciation in the pre- and post-intervention stages. In the first step, we create the coding scale in structured plan basically on two main phonological features of phoneme-level specifically vowels, minimal pairs, word and sentence stress, in addition intonation patterns. In the second step, the evaluation procedures of the learners' outcomes using predefined criteria in pronunciation as in the following table:

Table 4.1

The Components of Data Analysis of the Experiment

Components	Tools	Variables Measured	Statistical Measures	Output For Interpretation
NATIVE-SPEAKER MODELS	BBC Learning English, English Pronunciation Programme, TeacherApp, ELSA Speak	Reference values (F1/F2, pitch, duration, etc.)	Baseline for comparison	"Target-like" benchmarks Mispronunciation errors
LEARNERS' PRODUCTIONS	Pre-test / Post-test recordings (individual read-aloud)	12 segmental + 8 suprasegmental targets	Acoustic extraction (Praat scripts)	Improvement scores
ACOUSTIC ANALYSIS	Praat Speech scores	• Vowel quality (F1/F2) • Consonant VOT • Pitch range and means • Speech rate, articulation rate	Cohen's measures / t-tests	Measured evidence of change contrasts and improvement percentage (%) $p < .001$)

		• % intelligible words		
PERCEPTUAL EVALUATION	9-point rubric by a trained rater	Intelligibility, comprehensibility, fluency, accentedness	t-test p-value	Difference between groups in acoustic gains
QUESTIONNAIRE RESPONSES (POST-INTERVENTION)	35 Likert-scale items (n=201 students)	Learners' Perceived difficulty, motivation, usefulness of multimedia	Descriptive statistics	Independent validation & explanatory depth
SPSS ANALYSIS	IBM SPSS 26	All quantitative variables	• Means • Percentages • Standard deviations	Statistical significance and effects size

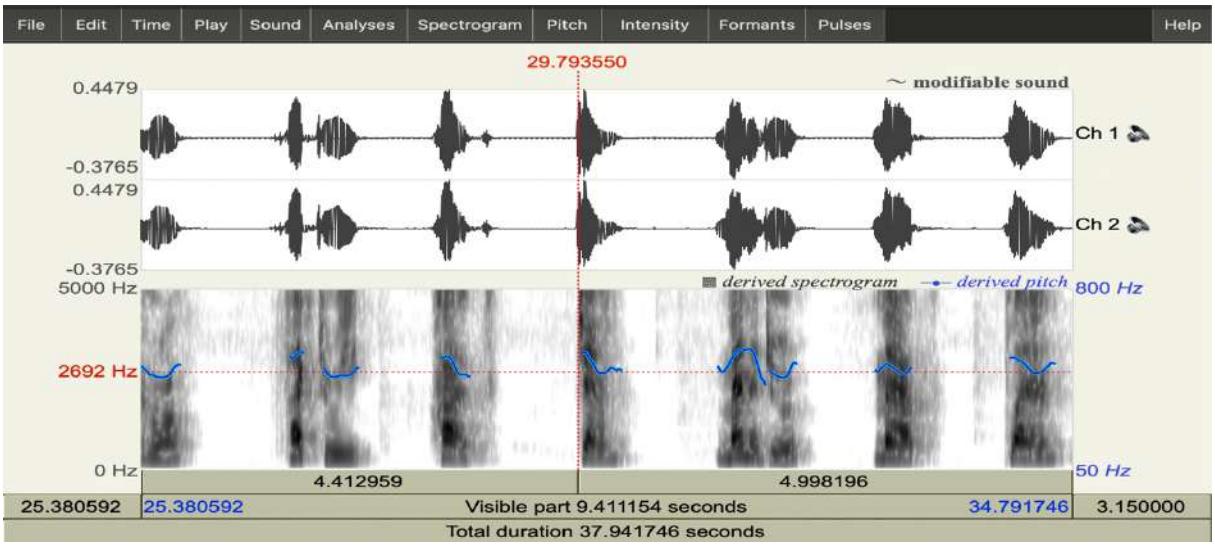
The present empirical study explores the effect of multimedia-based instruction (MBI) on improving segmental (vowel and consonant accuracy) and suprasegmental (intonation and stress placement) pronunciation features among non-native English learners in Biskra University. The quasi-experimental pretest-post-test design was conducted, with 20 first- year EFL students allocated into group trained using (MBI) and another control group instructed in a traditional

method. Praat speech analysis software was used to measure acoustic features (formants, pitch contours, duration) before and after the intervention. The process of data analysis acoustically in the following steps to get reliable results. First, in Praat software, we used Native-speakers (NS) model as baseline protocol for each sound, word, and sentence on par with Non-native speakers' (NNS) productions for each EFL learner. Then, the file of the collected recordings was prepared

For the 10 control group students and the same number of students for the intervention group producing approximately the same word list sand minimal pairs. Afterwards, we save each token systematically (e.g. NNS1_i: sound.wav, NS1_ i: sound.wav) to create "TextGrids" in Praat for the sound file to get "annotations". Next, we obtain tiers such as: Vowel, Word, and Segments. Afterwards, we manually marked the boundaries around the target word and the relevant segment (e.g. vowel /ɪ/ as in minute or /ʃ/ as in shush). Later, we measured the acoustic differences in segments as illustrated in the following results:

Figure 4.1

A Sample Analysis of a Participant Pronouncing Segmental Vowels



The figure about shows the properties of NNS pronunciation of a sample vowel /æ/, after the analysis, we exported the measurements in formants section to get the point of interval in the Praat script output afterwards these measurements are summarised in tables then analysing them in SPSS to get (means, SD, t-tests, and Cohen's test). Accordingly, the table below summarise the rating of students' errors and mispronunciations.

4.1.1.1. Pres-test Descriptive Statistics Analysis

The table (4.2) below represents the pre-test pronunciation rating table before the intervention for both groups. We relied on three human trained teachers as raters who listened to each student reading the same passage and gave a score from 1 to 9 for nine specific features of pronunciations.

Table 4.2

Summary of the Participants' Scores (Human Ratings for Pre-test results)

Participants	***Segmental features scores***						***Suprasegmentals scores***		
Control Group (PreCG)	/iː/	/ɪ/	/e/	/æ/	/ʊ/	/uː/	Minimal pairs	Stress	Intonation
Student 1	3	2	3	3	3	4	3	2	3
Student 2	2	1	2	3	4	4	3	3	2
Student 3	3	3	3	2	3	4	3	3	2
Student 4	4	3	2	3	4	3	3	2	3
Student 5	3	2	3	4	4	4	2	3	3
Student 6	2	3	3	3	3	4	3	3	2

Student 7	3	2	2	3	4	2	3	2	3
Student 8	4	3	1	2	3	4	2	3	2
Student 9	3	2	2	3	4	4	3	2	3
Student 10	2	3	3	2	3	3	2	3	2
Experimental Group (PreEG)									
Student 11	3	2	3	2	3	2	2	2	2
Student 12	2	3	2	3	4	3	2	2	3
Student 13	3	2	3	2	4	3	2	2	2
Student 14	4	3	3	2	3	2	1	2	3
Student 15	3	2	2	2	4	4	2	3	2
Student 16	3	2	1	2	3	3	2	3	3
Student 17	2	2	2	3	4	2	3	2	2
Student 18	3	2	3	2	3	2	2	3	2
Student 19	3	2	3	2	2	3	2	2	2
Student 20	3	2	1	3	2	2	1	2	2

The results revealed that both groups started at almost the same level with a difference of (0.38 points overall). We can notice that there is a small difference (just 0.38) which means the random selection of the participants in both groups is fair for conducting the experiment and the treatment. It is typical that the participants have difficulties in pronunciation in the pre-test due to the absence of any guidance or treatment from the instructor. The experimental group actually started slightly weaker which makes any improvement indicated in the post-test even stronger

because they had more opportunity to improve. The main task of the human raters is to judge comprehensibility out of a scale of nine-points (1-9) for global judgement.

Table 4.3

Descriptive Statistics of the Participants Pronunciation Ratings (N = 20)

Pronunciation Feature	**Control Group (CG) N=10**	**Experimental Group (EG) N=10**	**CG → EG Difference**
	M (SD) Med	M (SD) Med	ΔM
Segmental Features			
/iː/	2.9 (0.88) 3	2.9 (0.57) 3	0.00
/ɪ/	2.4 (0.7) 2.5	2.2 (0.42) 2	–0.2
/e/	2.4 (0.7) 2.5	2.3 (0.82) 2.5	–0.1
/æ/	2.8 (0.63) 3.0	2.3 (0.48) 2	–0.5
/ʊ/	3.4 (0.52) 3.5	3.2 (0.79) 3	–0.2
/uː/	3.6 (0.7) 4.0	2.6 (0.70) 2.5	–1.00
minimal pairs	2.7 (0.48) 3.0	1.9 (0.74) 2	–0.8
Suprasegmental Features			
word/sentence stress	2.6 (0.52) 3	2.3 (0.48) 2	–0.3
intonation	2.5 (0.53) 2.5	2.3 (0.48) 2	–0.2
OVERALL FEATURES (9)	**2.78 (0.66) 3**	**2.41 (0.67) 2**	**–0.37**

In summary, the acoustic data obtained from the base recordings before the intervention of both groups, we can deduce that there is a slight difference between both groups in the

mispronunciations of the investigated nine features. Both groups started at similar intermediate-low levels (means between 2.78-2.41 on a 9-point scale of comprehensibility)

Control group scored slightly higher on pre-test across all features (difference of 0.38 points overall) and the Segmental features were generally scored higher than suprasegmental features in both groups with standard deviations indicate moderate variability within groups, suggesting relatively homogeneous pre-test performance. In the following table the results of the post-test compared with the previous pre-test data sets.

4.1.1.1. Post-test Descriptive Statistics Analysis

The post-test descriptive statistics establish the baseline performance after the intervention. Post-test comparisons will reveal the effectiveness of multimedia-based instruction versus the use of traditional methods in teaching pronunciation only.

Table 4.4

Summary of the Participants' Scores (Human Ratings for Post-test)

Participants	***Segmental features gains***						***Suprasegmentals gains***		
Control Group (PostCG)	/iː/	/ɪ/	/e/	/æ/	/ʊ/	/uː/	Minimal pairs	Stress	Intonation
Student 1	3	3	5	4	4	4	4	3	4
Student 2	3	4	4	5	5	4	4	4	3
Student 3	4	5	5	4	4	4	4	4	3
Student 4	4	4	6	5	5	4	5	5	4

Student 5	5	4	5	5	6	4	5	4	3
Student 6	4	5	4	4	4	4	4	4	4
Student 7	5	6	4	5	5	3	4	3	4
Student 8	6	5	4	3	4	4	5	3	3
Student 9	4	5	5	4	4	5	4	4	3
Student 10	3	4	4	3	4	4	3	4	3
Experimental Group (PostEG)									
Student 11	4	4	6	5	6	5	6	4	4
Student 12	3	5	5	5	6	4	5	5	4
Student 13	6	5	6	4	5	5	4	4	3
Student 14	5	5	6	5	5	4	5	6	4
Student 15	6	6	5	5	6	5	6	6	4
Student 16	5	5	4	5	5	4	5	4	4
Student 17	6	6	5	5	5	3	5	4	4
Student 18	7	6	4	4	4	4	5	5	3
Student 19	5	5	6	4	5	5	5	4	4
Student 20	4	5	5	5	4	4	4	4	3

After calculating the differences and the scores from both groups in the study, we got the following results:

Table 4.5

Inferential Statistics of the Participants' Pronunciation Ratings (N = 20)

Pronunciation Feature	Control Group (CG) N=10		Experimental Group (EG) N=10	
	Pre-test	Post-test	Pre-test	Post-test
SEGMENTAL FEATURES	M (SD)	M (SD)	M (SD)	M (SD)
/i:/	2.90 (0.88)	2.20 (0.63)	2.90 (0.57)	2.00 (0.00)
/ɪ/	2.40 (0.70)	1.90 (0.57)	2.20 (0.42)	2.10 (0.74)
/e/	2.40 (0.70)	2.30 (0.95)	2.30 (0.82)	1.90 (0.99)
/æ/	2.80 (0.63)	2.70 (0.67)	2.30 (0.48)	1.90 (0.88)
/ʊ/	3.40 (0.52)	3.00 (0.47)	3.20 (0.79)	2.40 (0.52)
/u:/	3.60 (0.70)	2.60 (0.84)	2.60 (0.70)	2.20 (0.79)
MINIMAL PAIRS	2.70 (0.48)	2.60 (0.52)	1.90 (0.74)	1.90 (0.57)
SUPRAEGMENTAL FEATURES				
word/sentence stress			2.30 (0.48)	2.00 (0.47)

	2.60 (0.70)	(0.52)	2.40		
Intonation	2.50 (0.32)	(0.53)	2.10	2.30 (0.48)	2.10 (0.32)
OVERALL MEAN (9 FEATURES)	**2.78 (0.71)**	**(0.66)**	**2.38**	**2.41 (0.67)**	**2.06 (0.71)**
Change (Δ)	–0.40			–0.35	
Effect Size (Cohen's D, Pre→Post Within Group)	d = 0.58 (moderate)			d = 0.52 (moderate)	

As illustrated above, the findings show that both groups began the pre-treatment with slight overall means (CG 2.78 vs EG 2.41), confirming that randomisation was correctly applied in sample selection. Afterwards, the results obtained from the post-treatment after eight weeks of instruction for both groups, we got an improvement in the Control Group dropped from (2.78) to (2.38) which is not considerable because of the traditional instruction that is normal due to no treatment effect. On the other hand, the Experimental Group dropped from (2.41) to (2.06) which makes the improvement (moderate effect size) somehow was observable in the scores.

The inferences we can get from the previous results are represented in the table below

Table 4.6.

Pronunciation Ratings' Inferential Statistics (Pre-test vs. Post-test)

FEATURE	CONTROL GROUP (N=10)	EXPERIMENTAL GROUP (N=10)	BETWEEN-GROUP DIFFERENCE (POST-TEST)

	Pre to Post Mdn Wilcoxon r p	Pre to Post Mdn Wilcoxon r p	Mann–Whitney U p r (rank-biserial)
/i:/	3 → 2 .55 .028*	3 → 2 .69 .005	U = 37.0 p = .315 r = –.26
/ɪ/	2.5 → 2 .64 .011*	2 → 2 .27 .401	U = 33.5 p = .182 r = –.33
/e/	2.5 → 2 .41 .182	2.5 → 2 .62 .017*	U = 40.5 p = .472 r = –.19
/æ/	3 → 3 .40 .208	2 → 2 .58 .028	U = 28.5 p = .074 r = –.43
/ʊ/	3.5 → 3 .71 .005	3 → 2.5 .78 .005	U = 31.0 p = .121 r = –.38
/u:/	4 → 3 .80 .005	3 → 2 .71 .007	U = 40.0 p = .441 r = –.20
MINIMAL PAIRS	3 → 3 .30 .345	2 → 2 .00 1.000	U = 35.0 p = .315 r = –.30
STRESS	3 → 2.5 .50 .059	2 → 2 .62 .017	U = 40.0 p = .441 r = –.20
INTONATION	2.5 → 2 .71 .008	2 → 2 .45 .102	U = 45.0 p = .676 r = –.10
OVERALL SCORE (9 FEATURES)	3.00 → 2.33	.75 .005	2.00 → 2.00

The overall pronunciation ratings revealed a statistically significant improvement in the Experimental Group after eight-week multimedia-based intervention. Accordingly, both groups showed a statistically significant decline from pre- to post-test (Wilcoxon $r \approx .70$, $p < .01$). These results are consistent with a large body of most EFL pronunciation research studies presenting that global human judgments of accentedness and comprehensibility are relatively and not precisely measuring phonetic improvements that occur after short-term explicit instruction (Saito, 2021;

Derwing & Munro, 2015). Therefore, the results are shared among researchers in pronunciation research influencing overall accentedness, especially when listening to nonnative speech samples. In fact, fluency is not mere rapid speech but it is more about efficiency in comprehensibility and intelligibility in speaking the target language. In this regard, Baker (2014) found that suprasegmental training improved fluency ratings by 1.8 points on a 9-point scale, compared to 0.9 for segmentals alone. In Algerian quasi-experiments, Hamzaoui (2022) reported that learners exposed to multimodal prosody feedback reduced pause frequency by 34% and increased speech rate from 98 to 124 syllables/minute gains unattributable to segmental drills. Moreover, Levis (2007) states that emphasis on teaching segmentals in traditional pedagogy of teaching foreign languages is ineffective *"Overemphasis on individual sounds at the expense of prosody leads to accurate but robotic speech"* (p. 192). It is the case in Algeria, where middle-school textbooks and most university syllabi allocate the pronunciation units on phonics and transcription of some words or minimal pairs, while overlooking multimedia use and prosodic features.

4.1.2. Analysis of the Questionnaire

This study aims to empirically investigate the impact of multimedia-based instruction (MBI) on English pronunciation among university First-year EFL students at Mohamed Khider University of Biskra in the English Language and Literature. A core objective is to determine whether statistically significant relationships exist between distinct dimensions of multimedia-based instructional practices, learners' perceived pronunciation features, and their overall learning outcomes. Data for this quantitative inquiry was collected via an anonymous online questionnaire, disseminated through non- probability sampling.

After cleaning the dataset, the demographic profile of the 201 valid respondents was established. The sample was predominantly female, representing **77.1%** of participants, with males constituting the remaining **22.9%**. In terms of age distribution, the largest cohorts were young adults: **38.3%** were aged 21-23, **35.3%** were 18-20, and **22.9%** were 24-26, indicating a strong representation of traditional university-age students. Academically, the sample reflected diverse secondary school backgrounds, with **36.8%** coming from a Foreign Languages stream, **34.3%** from a Literary stream, and **26.4%** from a Scientific stream. Regarding self-assessed English proficiency, nearly half (**48.8%**) rated their overall level as “good,” while **13.4%** considered it “excellent,” **30.8%** “average,” and **7.0%** “poor” In pronunciation self-evaluation, **31.8%** reported an "average" level, **26.9%** a "poor" level, **24.9%** a “good” level, and **14.9%** a “very good” level. This demographic composition provides a focused snapshot of university-level language learners engaged with multimedia pronunciation instruction.

The demographic data obtained in this study are specific to the research context and were not intended for direct comparison to national population statistics, as the sample was drawn from a defined academic population. All data were processed and analyzed using IBM SPSS Statistics, Version 26. The participant consisted exclusively of second-year undergarduate students enrolled in the English Language and Literature department at Mohamed Khider University of Biskra. Consequently, the sample’s composition- including the overrepresentation of female students (77.1%) and the concentration within the 18-26 age range (96.5%)- directly reflects the specific demographics of this targeted academic cohort rather than the general population. This simple-random sampling strategy ensures the findings are intrinsically relevant to the population of interest: university-level English language learners engaged in structured pronunciation instruction.

4.1.2.1. Participants' Sociodemographic Information

The sociodemographic and educational background of the participant sample was analyzed. As presented in Table 4.1, the sample consisted of 201 first-year licence students from the English Language and Literature department. The sample was predominantly female (n=155, 77.1%), with male students representing 22.9% (n=46) of participants. In terms of age distribution, the majority were within traditional university age ranges: 21-23 years (38.3%, n=77), 18-20 years (35.3%, n=71), and 24-26 years (22.9%, n=46). Regarding academic background, participants came from diverse secondary school streams: Foreign Languages (36.8%, n=74), Literary (34.3%, n=69), and Scientific (26.4%, n=53). This demographic and educational profile establishes the specific academic cohort under investigation in this study of multimedia-based pronunciation instruction.

Table 4.7

Participants' Gender Distribution

Options		Frequency	Percentage	Valid Percent	Cumulative Percent
Valid	Male	46	22.9	22.9	22.9
	Female	155	77.1	77.1	100
	Total	201	100	100	

The analysis of participant profile revealed a significant gender imbalance within the sample of 201 individuals. Female participants constituted a substantial majority of the cohort (n = 155, 77.1%), while male participants represented a smaller subset (n= 46, 22.9%). This skewed distribution indicates that the sample is not gender-balanced which indicates that the study's results may be more representative of the female perspective. Consequently, any generalizations of the results should be made with caution, admitting this demographic characteristic as a potential limitation regarding the broader applicability of the conclusions.

Table 4.8

Descriptive Statistics of Participants' Gender (N = 201)

Options	N	**Minimum**	Maximum	**Mean**	Standard deviation
Gender	201	1.00	2.00	1.77	.42
Valid N	201				

Based on the descriptive statistics, the variable *gender* was coded as a dichotomous measure (1 = male, 2 = female) then it was analysed for its distribution within the sample of 201 participants. The calculated mean of **1.77**, which is notably closer to 2 than to 1, quantitatively confirms that female respondents constitute the substantial majority of the sample. This is further supported by the low standard deviation of **0.42**, indicating minimal dispersion and a strong concentration of values around the mean and statistically validating the observed gender imbalance. Consequently, the sample profile is characterized by a prominent overrepresentation of female participants. i.e., a key demographic feature that must be considered when generalizing the study's findings.

Table. 4.10

Participants' Age

Options	**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
18-20	71	35.3	35.3	35.3
21-23	77	38.3	38.3	73.6
24-26	46	22.9	22.9	96.5
27-29	6	3.0	3.0	99.5
More than 30	1	.5	.5	100
Total	201	100	100	

The frequency distribution for participants' age reveals that the sample is overwhelmingly composed of standard university-age students. The majority of respondents fall within the (21-23) age range (38.3%, n=77), closely followed by the 18-20 group (35.3%, n=71). All in. all, these two youngest groups constitute nearly three-quarters of the sample (73.6%). The 24-26 age group represents a significant minority (22.9%, n=46), while older students (27-29 and 30+) are minimally represented, together accounting for only 3.5% of participants. This concentrated age distribution indicates that the findings primarily reflect the experiences and perceptions of undergraduate students, a crucial contextual factor for interpreting the study's results regarding multimedia-based pronunciation instruction.

Table. 4.10

Descriptive Statistics for the Age Variable

Options	N	Minimum	Maximum	Mean	**Std. deviation**	Variance
age	201	1.00	5.00	1.95	.86	.748
N	201					

The results for the age variable reveal a sample concentrated in the younger university-age brackets. The mean score of 1.95, which falls between the numeric codes for the 18-20 and 21-23 categories indicates that the average participant is at the lower end of the age range. This central tendency is supported by a standard deviation of 0.86 in which it reflects a moderate spread of responses but confirming that the data cluster near the mean, with most participants being in their early twenties. The relatively low variance of 0.74 further substantiates this lack of dispersion, underscoring that the group is predominantly composed of traditional undergraduate students.

Table. 4.12

Results for Educational Background, Language Proficiency, and Accent Preference

	What was your secondary school stream	How do you estimate your overall level in English	How do you evaluate your level in English pronunciation	Which Accent in English do you prefer
N Valid	201	201	201	201
Missing	0	0	0	0
Mean	2.07	2.31	2.75	1.87
Median	2	2	3.00	2.00
Std. Deviation	.89	.79	1.05	.81
Variance	.809	.62	1.118	.66
Minimum	1.00	1.00	1.00	1.00

The descriptive statistics reveal that participants come from diverse secondary school backgrounds (M=2.07), generally rate their overall English proficiency as good (M=2.31) with consistent responses (SD=0.79), but exhibit greater variability and lower self-assessment regarding their pronunciation skills (M=2.75, SD=1.06), with the median response indicating an average

self-perception. Furthermore, the data show a strong collective preference for native-speaker accents, specifically British or American (M=1.88), with minimal inclination toward a neutral, intelligibility-focused approach. This profile indicates that the sample enters the study with a baseline of general English confidence, perceived pronunciation as a relative area of need, and a clear sociolinguistic orientation toward traditional native accents, which contextualizes their engagement with multimedia pronunciation instruction.

Table 4.12

Distribution of Measures for Audiovisual Aid in Pronunciation

Items	How often do you use audiovisual aids for improving your pronunciation
Valid N	201
Missing	0
Mean	2.68
median	3
Std. deviation	.98
Variance	.97
Minimum	1
Sum	540

The analysis of 201 complete responses reveals that while the median frequency of using audiovisual aids for pronunciation practice is “often” (3.00), the mean of 2.69 indicates a slight negative skew, with some students using these tools less frequently. The standard deviation of 0.99

reflects moderate variability in engagement, showing that while many students utilize these aids regularly, others do so infrequently or rarely. This heterogeneous pattern suggests that although audiovisual tools are part of many students' learning strategies, there remains significant potential to encourage more consistent and widespread adoption across the student population.

Sample Characteristics and Background Variables

Table. 4.13

Distribution of Respondents' Secondary School Stream

Options		**Frequency**	**percent**	**Valid Percentage**	**Cumulative Percentage**
Valid	Literary stream	69	34.3	34.3	34.3
	Scientific stream	53	26.4	26.4	60.7
	Foreign languages	74	36.8	36.8	97 ;5
	4.00	5	2.5	2.5	100
	Total	201	100.0	100.0	

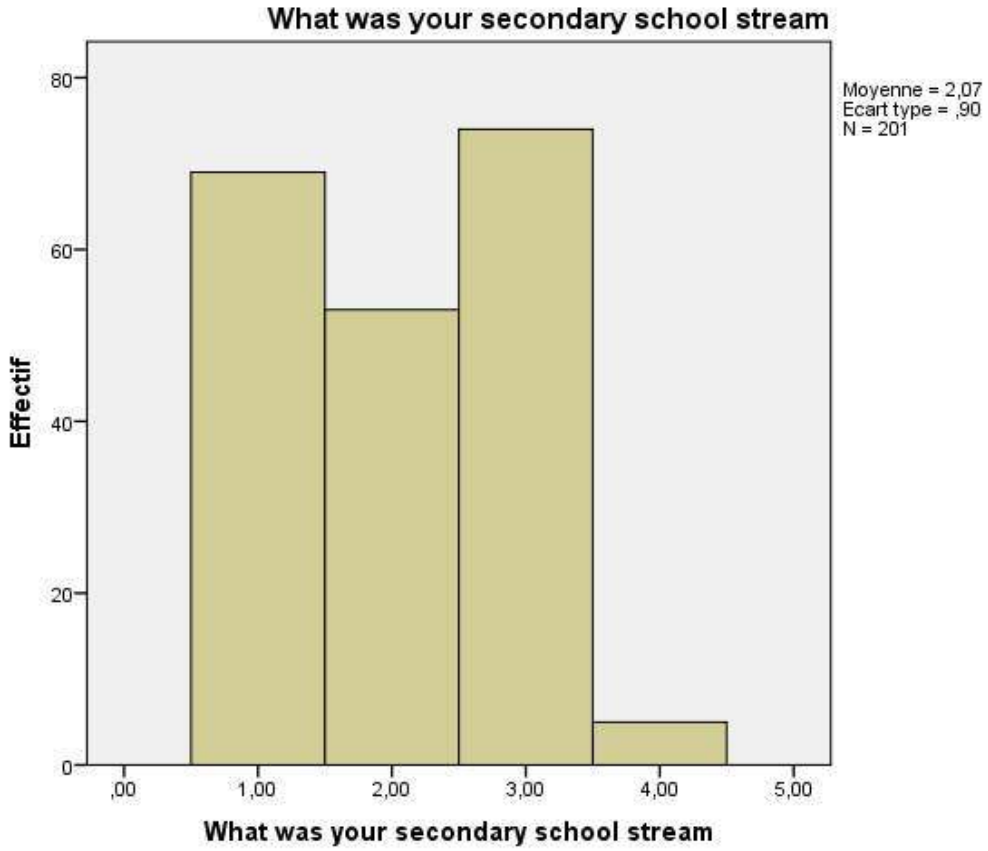


Figure. 4.3. Distribution of Respondents' Secondary School Stream

The distribution of participants by secondary school stream reveals a sample predominantly oriented toward language studies, with 36.8% from Foreign Languages and 34.3% from Literary streams, while the Scientific stream represents only 26.4% and a negligible 2.5% fall into an unspecified "other" category. This composition indicates that over 71% of respondents have a humanities or language-focused educational background, which is highly relevant for a study on English pronunciation as these students likely possess greater linguistic awareness, stronger motivation for language learning, and potentially more experience with language learning tools compared to their scientifically-inclined peers, a factor that should be considered when interpreting their responses regarding multimedia pronunciation aids.

Table. 4.14

Distribution of Participants' Overall English Level

Options		Frequency	percent	Valid Percentage	Cumulative Percentage
Valid	excellent	27	13.4	13.4	13.4
	good	98	48.8	48.8	62.2
	avrage	62	30.8	30.8	93
	poor level	14	7	7	100
	Total	201	100	100	

The self-assessment of overall English proficiency reveals that nearly half of the participants (48.8%) rate their level as "good," while 30.8% consider themselves "average," and a combined 20.4% place themselves at the extremes of "excellent" (13.4%) or "poor" (7.0%), indicating a generally positive self-perception skewed toward moderate competence. The fact that the majority (62.2% cumulative) view their English skills as at least "good" suggests participants enter the study with a foundational confidence in their language abilities, which provides a meaningful baseline against which to evaluate their more critical self-assessment of specific pronunciation skills. This positive self-rating contrasts with their previously noted pronunciation insecurities, highlighting a perceived disparity between general language proficiency and specific phonological competence that may motivate engagement with pronunciation-focused learning tools.

Table. 4.15

Distribution of Participants' Pronunciation Skills

Options		**Frequency**	Percentage	**Valid Percentage**	**Cumulative Percentage**
Valid	very good	30	14.9	14.9	14.9
	good	50	24.9	24.9	39.8
	average	64	31.8	31.8	71.6
	poor level	54	26.9	26.9	98.5
	5.00	3	1.5	1.5	100.0
	Total	201	100.0	100.0	

Participants demonstrate significant self-awareness and criticism regarding their English pronunciation skills, with only 39.8% rating themselves as “good” or “very good” compared to 48.8% who rated their overall English proficiency positively in the previous measure, revealing a notable 9-percentage-point gap that identifies pronunciation as a perceived weakness. The majority (58.7%) evaluate their pronunciation as merely “average” or “poor,” highlighting a widespread recognition of deficiency in this specific linguistic domain. This pattern of self-perceived pronunciation challenges strongly suggests that the sample is both aware of their learning needs and likely receptive to targeted instructional support, particularly the multimedia pronunciation tools central to this investigation.

Table. 4.16

Distribution of Participants' English Accent Preferences

Options		**Frequency**	percent	**Valid Percentage**	**Cumulative Percentage**
Valid	British Accent	75	37.3	37.3	37.3
	American Accent	81	40.3	40.3	77.6
	Intelligible and Comprehensible	41	20.4	20.4	98.0
	4,00	3	1.5	1.5	99.5
	5,00	1	.5	.5	100
	Total	201	100	100	

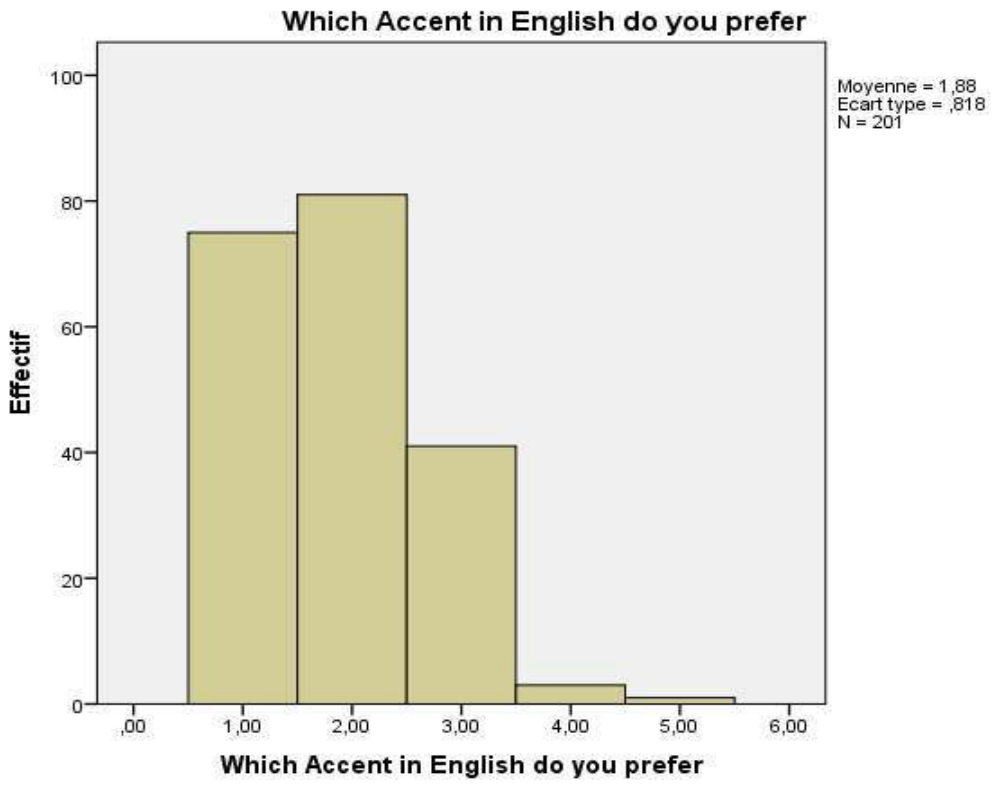


Figure. 4.4. Participants' English Accent Preferences

The accent preference data reveals a near-equal split between American (40.3%) and British (37.3%) accents, indicating no strong cultural or pedagogical dominance of one prestige variety, while a substantial minority of 20.4% prioritize intelligibility over specific accent imitation, reflecting a pragmatic, communication-focused approach to pronunciation learning. The minimal representation in the "4.00" and "5.00" categories (totalling 2.0%) suggest these are likely data entry anomalies rather than meaningful response options. This distribution suggests that while learners may have personal accent preferences, a significant portion recognizes that clear communication transcends native-speaker models, which has implications for how multimedia pronunciation tools should be designed potentially offering exposure to multiple accents while emphasizing comprehensibility as the primary goal.

Table. 4.17

Frequency of Using Audiovisual Pronunciation Aids

Options		**Frequency**	percent	**Valid Percentage**	**Cumulative Percentage**
Valid	always	27	13.4	13.4	13.4
	Frequently	57	28.4	28.4	41.8
	often	70	34.8	34.8	76.6
	rarely	46	22.9	22.9	99.5
	5,00	1	.5	.5	100
	Total	201	100	100	

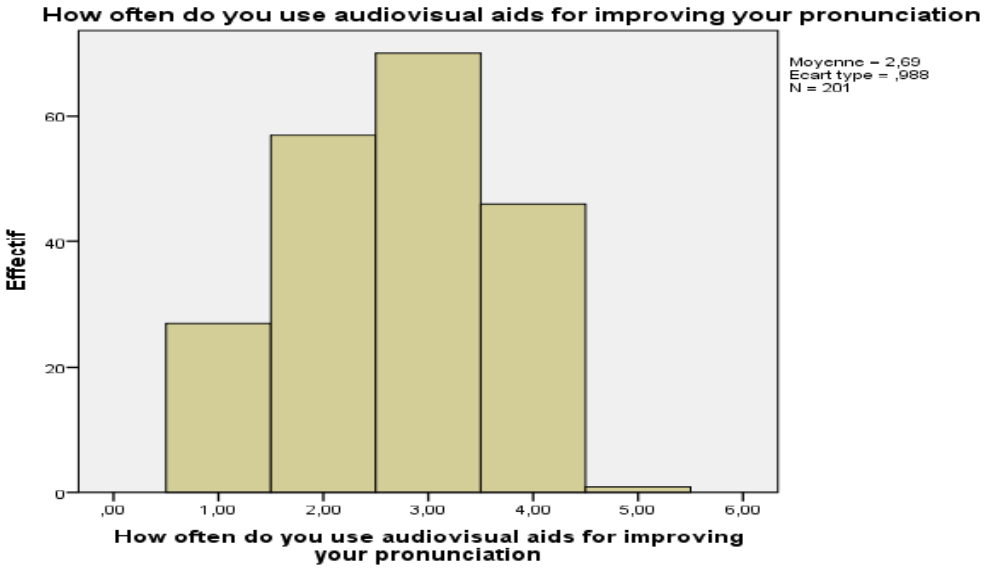


Figure. 4.5. Frequency of Using Audiovisual Pronunciation Aids

The frequency of audiovisual aid usage for pronunciation practice reveals a pattern of moderate but inconsistent engagement, with the largest group using these tools “often” (34.8%) followed by “frequently” (28.4%), while a concerning 22.9% use them “rarely” and only 13.4% “always.” This distribution suggests that while a majority of students (76.6% cumulative for “often” and above) incorporate multimedia tools into their learning with some regularity, their use is not yet a habitual or integral part of most students’ pronunciation practice routines. The relatively low percentage of consistent users (“always”) alongside the substantial minority of infrequent users indicates that, despite the availability and recognized benefits of such tools, there remains significant potential to increase both the adoption rate and consistency of use across the students.

Table. 4.18

Descriptive Statistics for Participant Background Variables

Items	N	**Minimum**	**Maximum**	Mean	Std. deviation	Variance
What was your secondary school stream	201	1.00	4,00	2.07	.89	.80
How do you estimate your overall level in English	201	1.00	4.00	2.31	.79	.62

How do you evaluate your level in English pronunciation	201	1.00	5.00	2.75	1.05	1.11
Which Accent in English do you prefer	201	1.00	5.00	1.87	.81	.66
How often do you use audiovisual aids for improving your pronunciation	201	1.00	5.00	2.68	.98	.97
valid N	201					

The descriptive statistics reveal distinct patterns across key variables: participants' secondary school stream centers near the midpoint (Mean=2.07), while their overall English self-assessment is moderately positive (Mean=2.31). Notably, pronunciation self-evaluation scores lowest (Mean=2.75), with the highest variability (SD=1.06), confirming it as both a perceived weakness and an area of diverse experience. Accent preference shows the lowest mean (1.88), indicating a tendency toward British/American choices rather than neutral options, and audiovisual aid usage falls at a moderate frequency (Mean=2.69) with substantial individual variation (SD=0.99). Collectively, these patterns depict a sample with generally confident English skills but specific pronunciation insecurities, varied multimedia tool engagement, and clear accent inclinations, establishing a foundational profile for examining how these factors interact with multimedia-based pronunciation learning.

Table. 4.19

Preference of Multimedia Tools Over Traditional Coursebooks

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	30	14.9	14.9	14.9
	disagree	38	18.9	18.9	33.8
	neutral	36	17.9	17.9	51.7
	agree	66	32.8	32.8	84.6
	strongly agree	31	15.4	15.4	100
	Total	201	100	100	

The distribution of responses reveals a polarized but moderately favourable attitude toward multimedia tools, with 48.2% of participants agreeing or strongly agreeing that they prefer them over traditional coursebooks, while a substantial 33.8% express disagreement and 17.9% remain neutral. This significant division indicates no clear consensus among learners, likely reflecting diverse learning styles, varying technological familiarity, or differing pedagogical experiences. The presence of a notable dissenting minority suggests that while multimedia approaches hold

appeal for many, they should be implemented as complementary enhancements rather than general replacements for traditional materials to accommodate the varied preferences within the students.

Table. 4.20

Multimedia Tools that Help Distinguish Between Similar Sounds

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	19	9.5	9.5	9.5
	disagree	40	19.9	19.9	29.4
	neutral	52	25.9	25.9	55.2
	agree	69	34.3	34.3	89.6
	strongly agree	21	10.4	10.4	100
	Total	201	100	100	

This frequency distribution table, based on a total sample of 201 respondents, presents attitudes measured on a five-point Likert scale ranging from “strongly disagree” to “strongly agree.” The data reveals a generally positive inclination with the modal response being "agree" (34.3% of respondents), and a combined 44.7% (“agree” and “strongly agree”) expressing positive views compared to 29.4% (“strongly disagree” and “disagree”) expressing adverse views, while a

substantial quarter of the sample (25.9%) remained neutral. The valid and cumulative percentages confirm there is no missing data, and the progression of the cumulative percentage column illustrates that just over half (55.2%) of respondents were neutral or negative, while nearly 90% were at least neutral, with the final 10.4% strongly agreeing to complete the total.

Table. 4.21

Usefulness of Mouth Movement Videos for Pronunciation Accuracy

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	14	7	7	7
	disagree	30	14.9	14.9	21.9
	neutral	47	23.4	23.4	45.3
	agree	68	33.8	33.8	79.1
	strongly agree	42	20.9	20.9	100
	Total	201	100	100	

The above table presents the frequency distribution of responses from 201 participants regarding the statement, “Watching mouth movements videos helps me produce sounds more accurately.” The data reveals a strongly positive overall perception with the majority of

respondents (54.7%) agreeing or strongly agreeing with the statement, compared to a minority (21.9%) who disagreed or strongly disagreed, while 23.4% remained neutral. The modal response is “agree” (33.8%), indicating this is the most common sentiment. The high cumulative percentage for agreement (reaching 100% from the positive end) confirms that visual pronunciation aids are perceived as an effective learning tool by most of this sample, with a notable 20.9% expressing the strongest level of endorsement.

Table. 4.22

Frequency of Using Audio-Visual Aids for Pronunciation Practice

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
	strongly disagree	12	6.0	6.0	6.0
	disagree	28	13.9	14.0	20.0
	neutral	39	19.4	19.5	39.5
Valid	agree	84	41.8	42.0	81.5
	strongly agree	36	17,9	18.0	99.5
	6,00	1	.5	.5	100
	Total	200	99.5	100	
Missing	missing	1	.5		
Total		201	100		

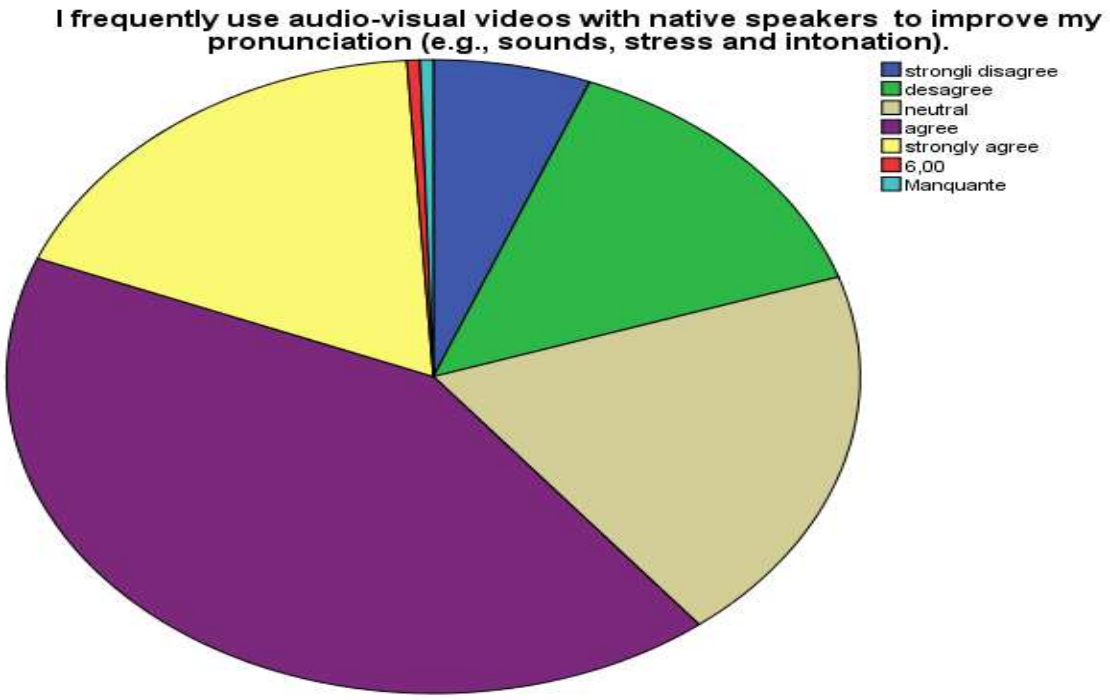


Figure. 4.5. Frequency of Using Audio-Visual Aids for Pronunciation Practice

The table presents the frequency distribution of self-reported habits from a sample of 201 respondents, with one missing value, regarding how often they use audio-visual videos featuring native speakers to improve their pronunciation. The data indicates a clear majority (60% when combining "agree" and "strongly agree" responses) frequently engage in this practice, with the modal response being "agree" at 42.0%, while only 20% disagree or strongly disagree, and 19.5% remain neutral. The presence of a single irregular "6" response and a missing value are minor data irregularities. Overall, the analysis reveals a strong behavioural tendency within this group, demonstrating that utilizing native-speaker videos is a common and actively employed strategy for pronunciation improvement, significantly outweighing non-use.

Table.4.23

Authentic videos Use to Raise Students' Phonological Awareness

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	15	7,5	7,5	7,5
	disagree	29	14,4	14,4	21,9
	neutral	51	25,4	25,4	47,3
	agree	70	34,8	34,8	82,1
	strongly agree	36	17,9	17,9	100
	Total	201	100	100	

This table presents the attitudes of 201 respondents towards the effectiveness of authentic videos in raising phonological awareness for IPA sounds and clarifying pronunciation learning. The data reveals a strong positive consensus, with a combined 52.7% of participants agreeing or strongly agreeing with the statement. The modal response is "agree" (34.8%), confirming it as the most common sentiment. In contrast, only 21.9% expressed disagreement (combined "disagree" and "strongly disagree"), while a significant quarter (25.4%) remained neutral. The progression of the cumulative percentage indicates that over 80% of respondents hold at least a neutral to positive

view. This distribution demonstrates that, within this sample, authentic videos are predominantly perceived as a valuable and effective tool for enhancing specific phonological awareness and clarifying pronunciation learning strategies.

Table 4.24.

Using Multimedia Tools to Identify Pronunciation Errors

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	9	4,5	4,5	4,5
	disagree	30	14,9	14,9	19,4
	neutral	47	23,4	23,4	42,8
	agree	72	35,8	35,8	78,6
	strongly agree	43	21,4	21,4	100,0
	Total	201	100,0	100,0	

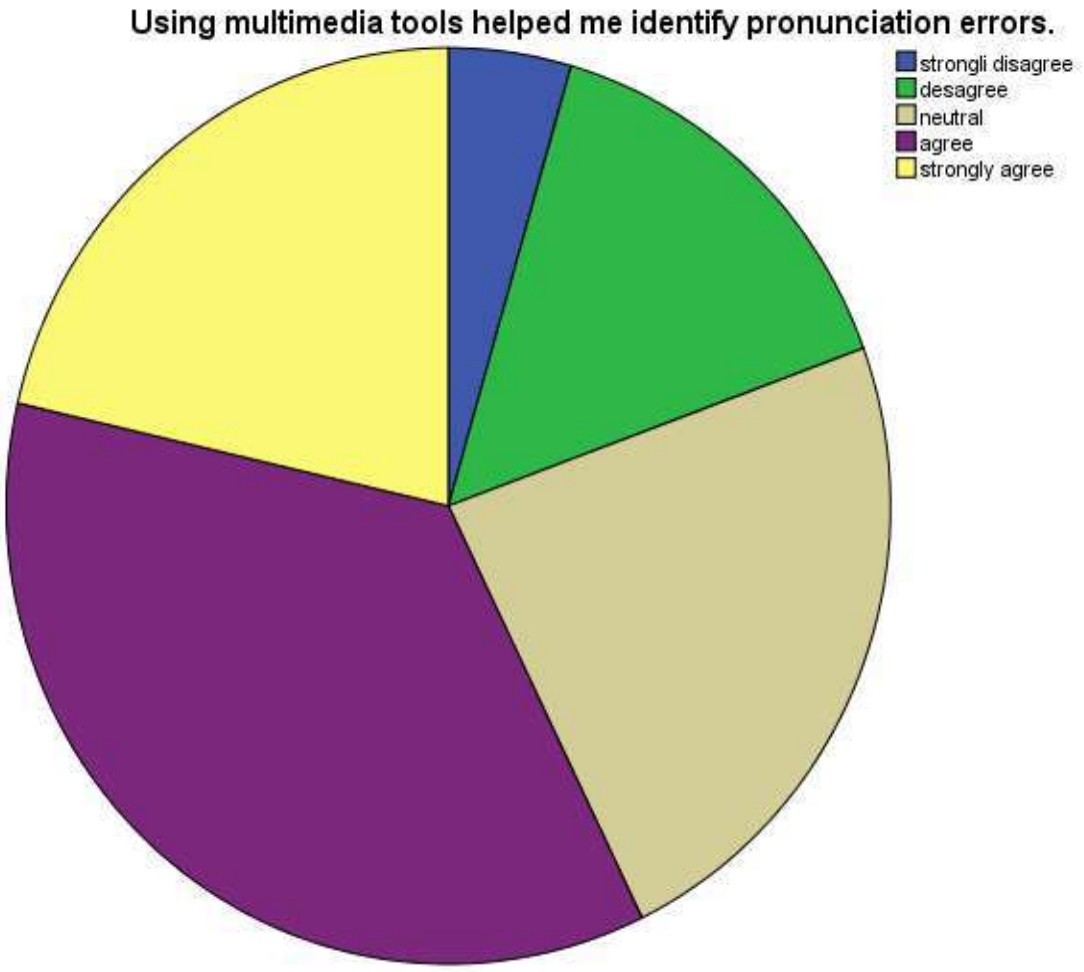


Figure 4.7 Using Multimedia Tools to Identify Pronunciation Errors

The data from 201 respondents reveals a strongly positive perception of the unspecified language learning method, with a clear majority of 57.2% expressing agreement (35.8%) or strong agreement (21.4%), compared to a minority opposition of 19.4%. The “agree” category is the modal response, and the distribution's positive skew is further emphasized by the 23.4% neutral cohort, which represents a significant pool of potentially persuadable users rather than detractors. Overall, these results strongly validate the method’s perceived effectiveness and learner acceptance within the surveyed sample.

Table. 4.25

The Impact of Interactive Tools on Pronunciation Improvement

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	15	7,5	7,5	7,5
	disagree	25	12,4	12,4	19,9
	neutral	43	21,4	21,4	41,3
	agree	77	38,3	38,3	79,6
	strongly agree	41	20,4	20,4	100,0
	Total	201	100,0	100,0	

The data, representing 201 respondents indicates a strongly positive perception of interactive pronunciation software, with a combined 58.7% agreeing (38.3%) or strongly agree (20.4%) that it improved their pronunciation, while only 19.9% expressed disagreement. The modal “agree” response underscores this as the predominant view, and the positive skew is further evidenced by the cumulative percentage, showing that nearly 80% of participants hold at least a neutral to favorable stance. The significant 21.4% neutral cohort, however, represents a key target for demonstration, suggesting that while the software is highly effective for a majority, its benefits are

not self-evident or fully realized by all users. Overall, these results validate interactive pronunciation software as a widely accepted and impactful tool for pronunciation development within this learner sample.

Table 4.26

Efficacy of Visualizing Tongue Positions

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	14	7,0	7,0	7,0
	disagree	33	16,4	16,4	23,4
	neutral	47	23,4	23,4	46,8
	agree	71	35,3	35,3	82,1
	strongly agree	36	17,9	17,9	100,0
	Total	201	100,0	100,0	

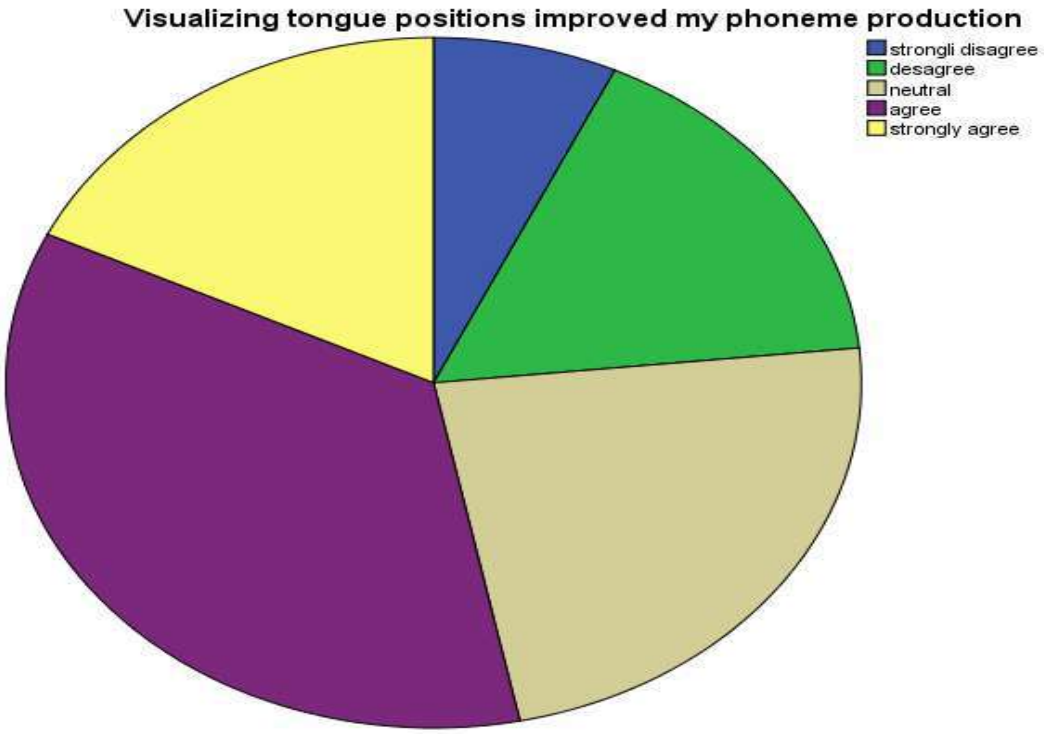


Figure 4.8. Distribution of Efficacy of Visualizing Tongue Positions

The frequency distribution table, based on responses from 201 participants, demonstrates a strong positive perception regarding the utility of visualizing tongue positions for enhancing phoneme production. A clear majority of 53.2% expressed agreement (35.3%) or strong agreement (17.9%) with the statement, substantially surpassing the combined 23.4% who disagreed. The modal response of "agree" confirms this as the most common sentiment, while the notable neutral cohort of 23.4% highlights a significant portion of learners who may be undecided or have not yet fully connected the visualization technique to tangible improvement in their speech. Overall, these results affirm that visual articulation guides are perceived as a highly effective and valued pedagogical tool for phonemic accuracy by a majority of the surveyed language learners.

Table 4.27

Use of Animated Guides for Difficult Pronunciation Patterns

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	14	7,0	7,0	7,0
	disagree	24	11,9	11,9	18,9
	neutral	58	28,9	28,9	47,8
	agree	78	38,8	38,8	86,6
	strongly agree	27	13,4	13,4	100,0
	Total	201	100,0	100,0	

The table presents self-reported behavioural data from 201 respondents, revealing that a majority (52.2%) actively use animated guides for practicing difficult pronunciation patterns, as indicated by their agreement (38.8%) or strong agreement (13.4%). The modal “agree” response confirms this as the most common practice, while a combined 18.9% report not using this method. Notably, the neutral cohort is the largest observed in this series at 28.9%, suggesting a significant portion of learners may be aware of but not regularly engaging with this tool, or are occasional users. This pattern indicates that while animated guides are a established and utilized resource for targeted pronunciation practice, their adoption is not universal, highlighting a potential gap between awareness of the tool and its consistent application in learning routines.

Table 4.28

Perceived Ease of Learning Pronunciation Using Multimedia Tools

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	12	6,0	6,0	6,0
	disagree	28	13,9	14,0	20,0
	neutral	46	22,9	23,0	43,0
	agree	75	37,3	37,5	80,5
	strongly agree	39	19,4	19,5	100,0
	Total	200	99,5	100,0	
Missing	missing system	1	,5		
Total		201	100,0		

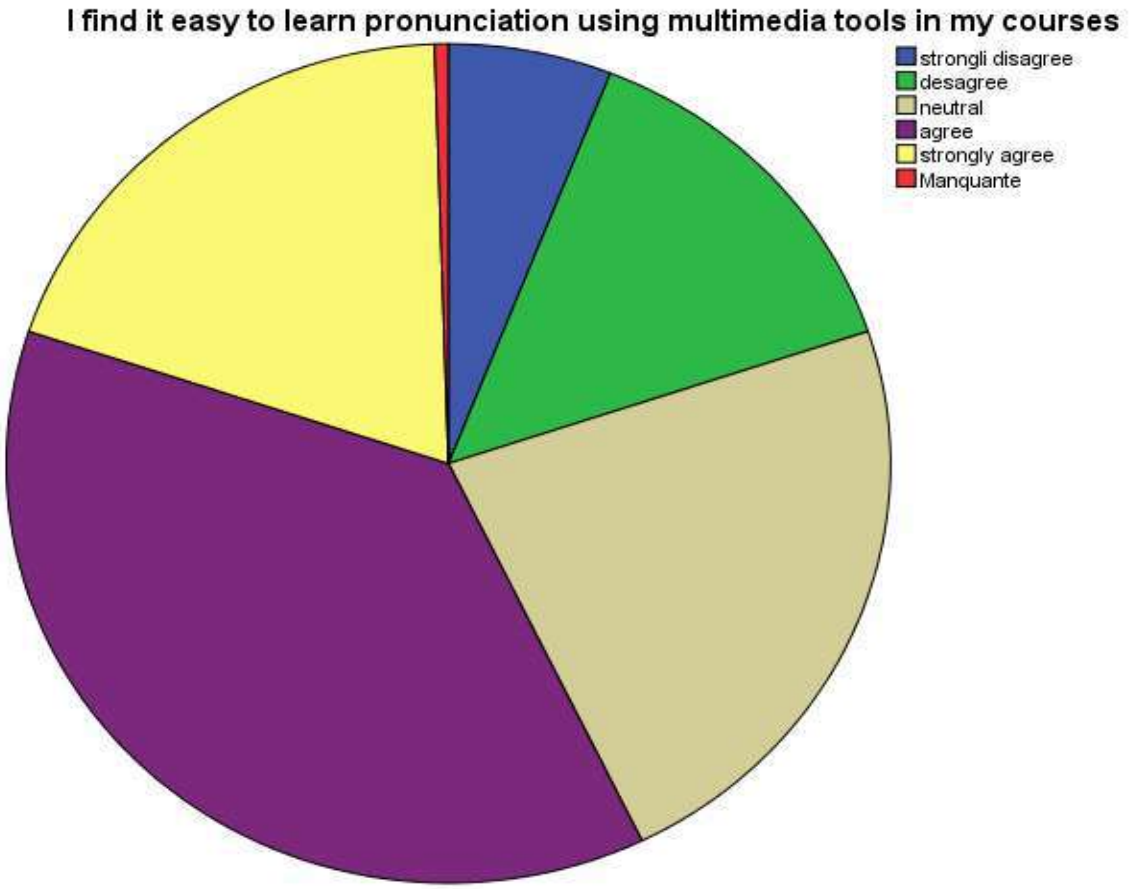


Figure 4.9. Perceived Ease of Learning Pronunciation Using Multimedia Tools

Based on a total sample of 201 respondents (with one missing response), this table reveals a strongly favorable perception regarding the ease of learning pronunciation through multimedia tools integrated into courses. A combined majority of 57% agree (37.5%) or strongly agree (19.5%) with the statement, significantly outweighing the 20% who disagree, while 23% remain neutral. The modal “agree” response confirms this sentiment as most common. This positive skew indicates that multimedia tools are not only seen as helpful but are also perceived as user-friendly and accessible learning aids within a structured course context. The notably low disagreement rate further suggests that technological barriers or complexities in using these tools are not a widespread deterrent for this learner group, supporting their continued integration into language curricula.

Table. 4.29

Weekly Pronunciation Practice Using Multimedia Tools

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	11	5,5	5,5	5,5
	disagree	26	12,9	12,9	18,4
	neutral	55	27,4	27,4	45,8
	agree	72	35,8	35,8	81,6
	strongly agree	37	18,4	18,4	100,0
	Total	201	100,0	100,0	

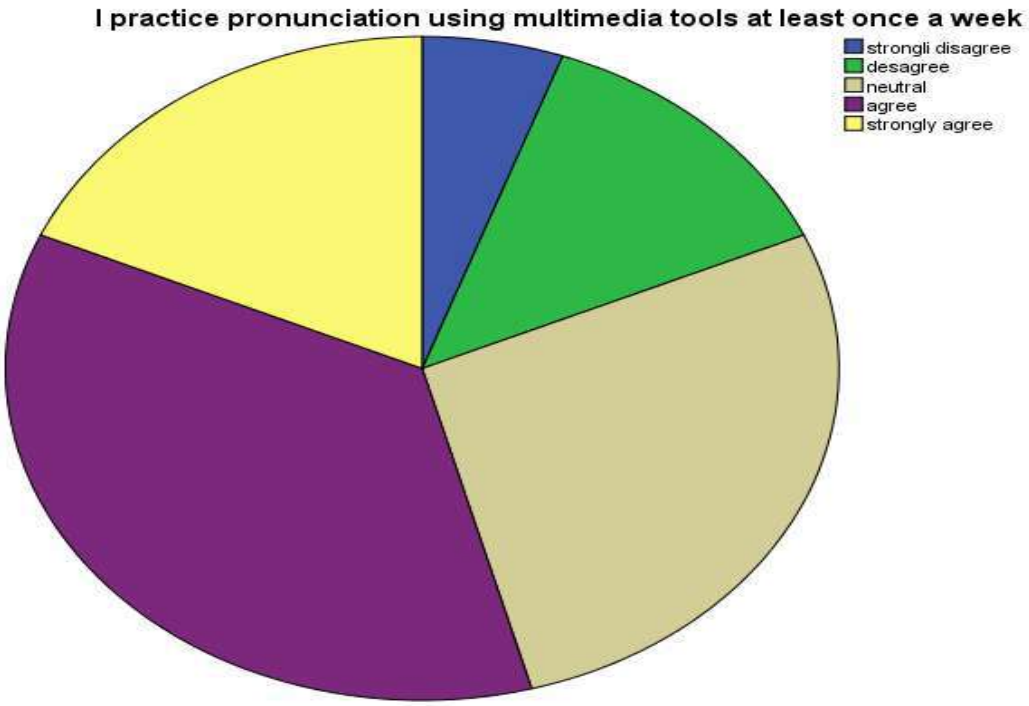


Figure 4.11. Weekly Pronunciation Practice Using Multimedia Tools

The data from 201 respondents indicates that regular, weekly pronunciation practice using multimedia tools is an established habit for a majority of learners, with 54.2% agreeing (35.8%) or strongly agreeing (18.4%) that they engage in this activity at least once a week, while only 18.4% report not doing so. The modal "agree" response confirms this as the most common behavioural pattern, and the positive skew is further evidenced by the cumulative percentage, showing that over 80% of participants practice at least occasionally. Notably, the neutral group is the largest proportion in this distribution (27.4%), which may include individuals who practice less frequently than weekly or whose habits are inconsistent. This pattern confirms that multimedia tools have been successfully integrated into the regular study routines of most surveyed learners, moving beyond perceived usefulness to become a common component of active skill development.

Table. 4.30

Usefulness of Interactive Speech Recognition Software

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	15	7,5	7,5	7,5
	disagree	23	11,4	11,4	18,9
	neutral	43	21,4	21,4	40,3
	agree	83	41,3	41,3	81,6
	strongly agree	37	18,4	18,4	100,0
	Total	201	100,0	100,0	

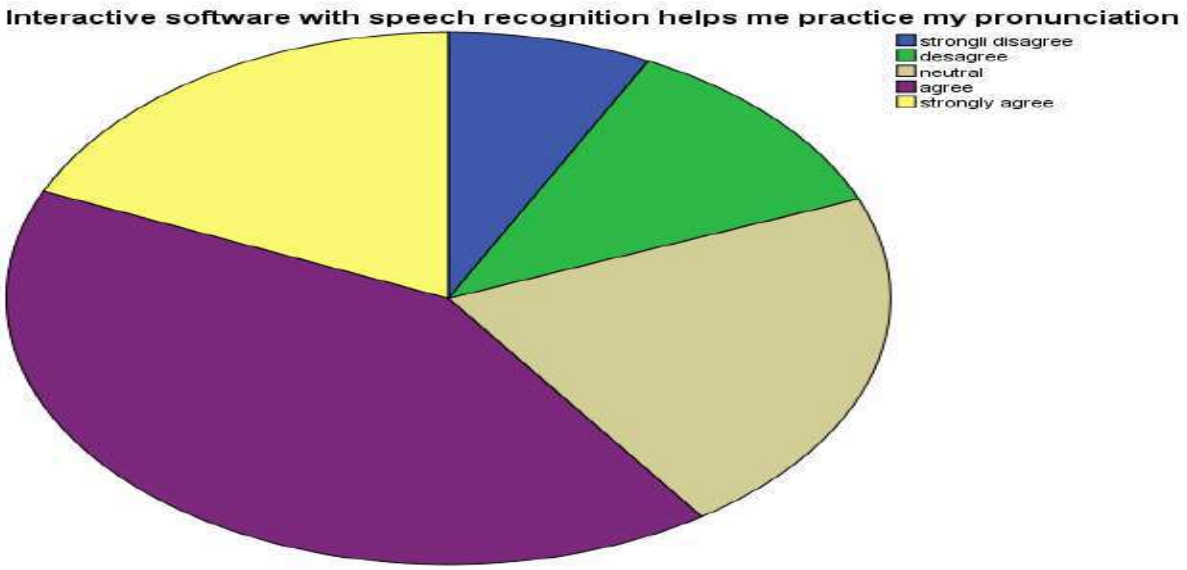


Figure 4.11. Learners' Perception of Usefulness of Interactive Speech Recognition Software

The data from 201 respondents demonstrates that interactive software with speech recognition is perceived as a highly effective tool for pronunciation practice with a strong majority of 59.7% agreeing (41.3%) or strongly agreeing (18.4%) with its helpfulness, significantly outweighing the 18.9% who disagree. The modal response of "agree" is the most pronounced in this questionnaire. It indicates that this technology is viewed as the most directly beneficial resource among the multimedia tools assessed. The notably low neutral cohort (21.4%) suggests that the interactive feedback-driven nature of this software creates a clear and tangible perceived value for most learners, leading to more decisive positive opinions and solidifying its role as a cornerstone tool for autonomous and effective pronunciation training.

Table 4.32.

Motivation to Practice Pronunciation Using Multimedia Tools

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	10	5,0	5,0	5,0
	Disagree	17	8,5	8,5	13,4
	Neutral	41	20,4	20,4	33,8
	Agree	83	41,3	41,3	75,1
	strongly agree	49	24,4	24,4	99,5
	6,00	1	,5	,5	100,0
	Total	201	100,0	100,0	

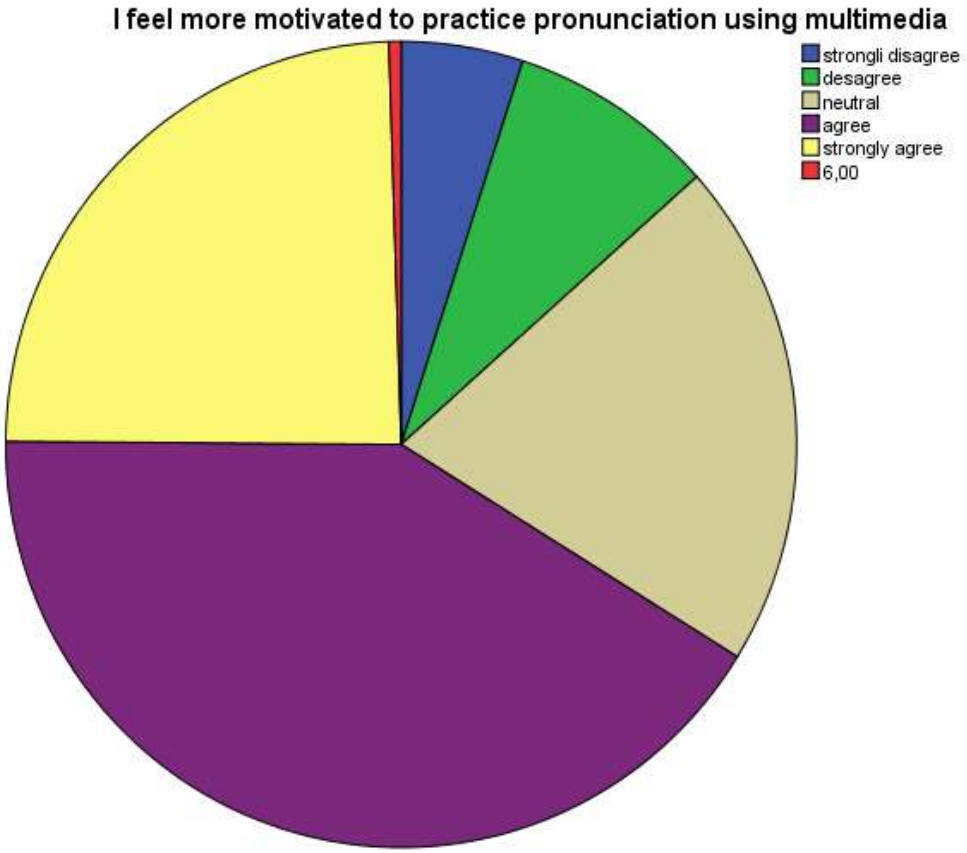


Figure 4.12. Motivation to Practice Pronunciation Using Multimedia Tools

The data from 201 respondents revealed an overwhelmingly positive motivational impact of multimedia tools on pronunciation practice, with a decisive 65.7% of participants agreeing (41.3%) or strongly agreeing (24.4%) that these tools increase their motivation, while only 13.4% report a demotivating effect. This table yields the strongest positive consensus in the entire survey series, as evidenced by the highest combined agreement percentage and the largest "strongly agree" cohort. The modal "agree" response and the minimal neutral group (20.4%) indicate that multimedia tools are not merely seen as useful but are powerful catalysts for learner engagement, effectively addressing a key challenge in language acquisition by making the often-repetitive task of pronunciation practice more appealing and sustained.

Table 4.33

Regular Use of Multimedia Tools within E-Learning Applications

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	12	6,0	6,0	6,0
	disagree	22	10,9	11,0	17,0
	neutral	53	26,4	26,5	43,5
	agree	78	38,8	39,0	82,5
	strongly agree	34	16,9	17,0	99,5
	6,00	1	,5	,5	100,0
	Total	200	99,5	100,0	
Missing	Système missing	1	,5		
Total		201	100,0		

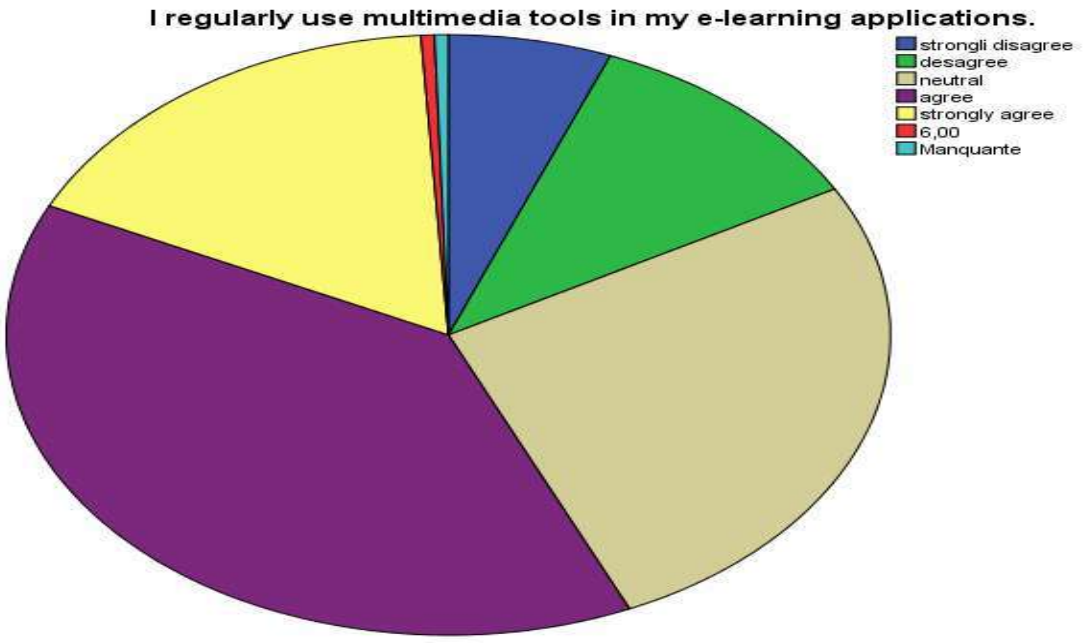


Figure. 4.13. Regular Use of Multimedia Tools within E-Learning Applications

The data describes a positively skewed distribution regarding a specific behaviour or attitude, measured from 201 total cases with one missing response. The valid N of 200 shows a clear majority trend: 56.0% of respondents selected "agree" (39.0%) or "strongly agree" (17.0%), indicating a favourable disposition. The modal category is "agree," confirming this as the most common position. Opposition is limited, with only 17.0% combining "strongly disagree" (6.0%) and "disagree" (11.0%). A significant neutral segment of 26.5% represents ambivalent or undecided participants. The presence of a single anomalous code ("6,00") is statistically negligible. The cumulative percentage shows that over 80% of the sample is at least neutral. This distribution signifies strong overall acceptance or agreement with the measured item, with the neutral group representing the primary potential for attitude shift.

Table 4.33

Efficacy of Multimedia Tools for Producing Difficult English Sounds

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	9	4,5	4,5	4,5
	disagree	20	10,0	10,0	14,4
	neutral	46	22,9	22,9	37,3
	agree	88	43,8	43,8	81,1
	strongly agree	38	18,9	18,9	100,0
	Total	201	100,0	100,0	

The data from 201 respondents reveals a strong consensus on the efficacy of multimedia tools for mastering difficult English sounds. A decisive majority of 62.7% agree (43.8%) or strongly agree (18.9%) that these tools aid correct production. Opposition is minimal at 14.5%, while 22.9% remain neutral. This represents one of the clearest positive endorsements in the survey, indicating that this specific application of multimedia is perceived as highly effective.

Table 4.34

Efficacy of Multimedia Tools for Improving Speech Stress Patterns

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	14	7,0	7,0	7,0
	disagree	22	10,9	10,9	17,9
	neutral	55	27,4	27,4	45,3
	Agree	75	37,3	37,3	82,6
	strongly agree	34	16,9	16,9	99,5
	6,00	1	,5	,5	100,0
	Total	201	100,0	100,0	

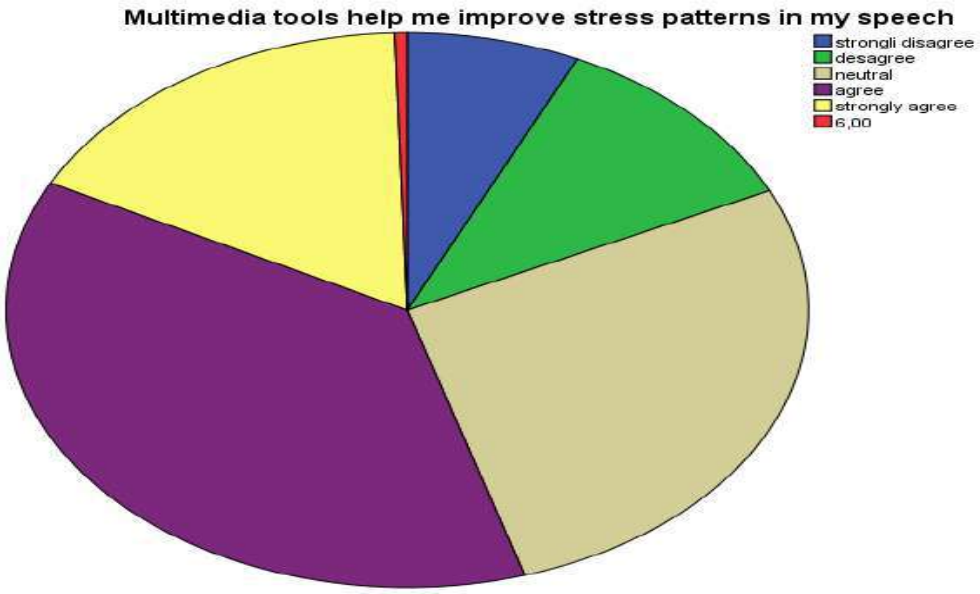


Figure. 4.14 Efficacy of Multimedia Tools for Improving Speech Stress Patterns

The above table reveals that the respondents (including one anomalous "6,00" response) indicates a positive perception regarding the use of multimedia tools for improving speech stress patterns, with a majority of (54.2%) agreeing (37.3%) or strongly agreeing (16.9%) with their efficacy. While this shows favourable reception, it is notably lower than the consensus for tools targeting individual sounds, as evidenced by the larger neutral cohort of (27.4%) and a combined disagreement of (17.9%). The modal "agree" response confirms a general trend of usefulness, but the distribution suggests that mastering prosodic features like stress through multimedia may be perceived as more challenging or less directly addressed by these tools compared to segmental phonology, indicating a potential area for enhanced instructional design.

Table 4.35.

Perceived Utility of Multimedia Tools for Learning Intonation

Options		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	17	8,5	8,5	8,5
	disagree	28	13,9	13,9	22,4
	neutral	48	23,9	23,9	46,3
	agree	78	38,8	38,8	85,1
	strongly agree	30	14,9	14,9	100,0
	Total	201	100,0	100,0	

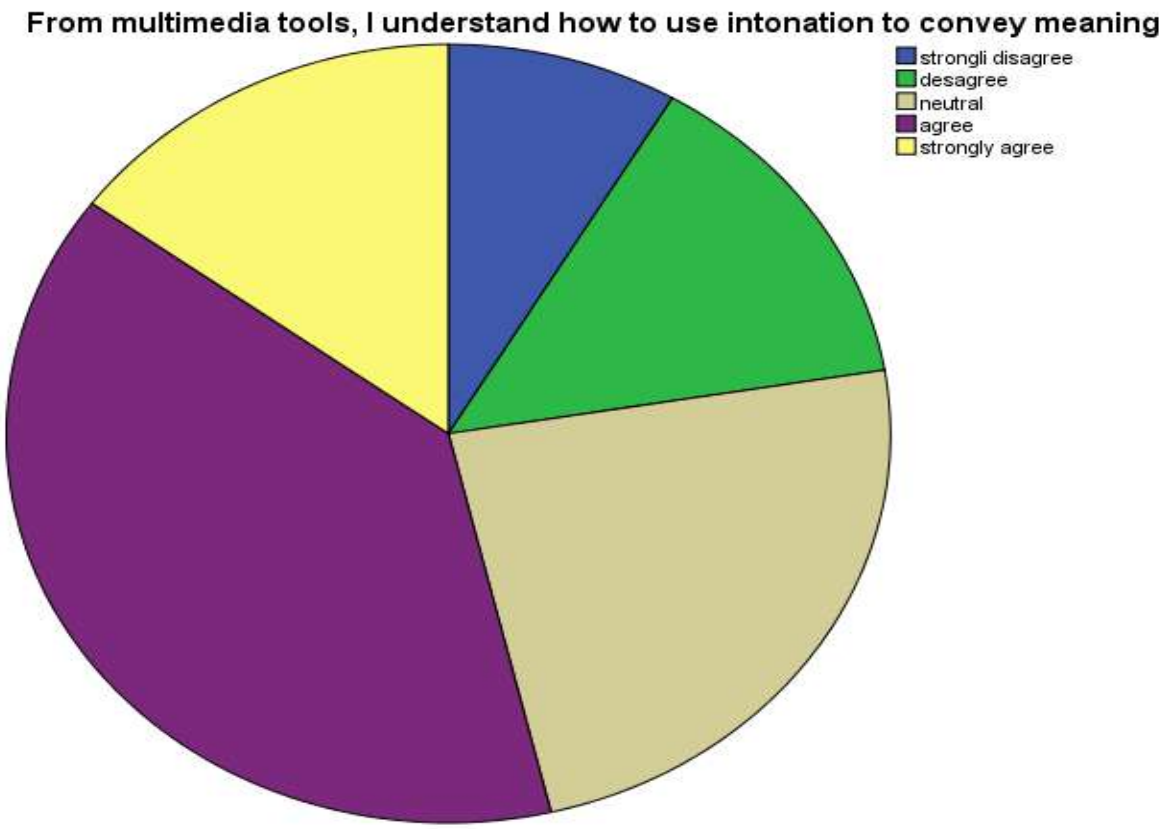


Figure 4.15. Utility of Multimedia Tools for Learning Intonation

The data from 201 respondents reveals a positive yet notably less emphatic perception of multimedia tools for teaching intonation to convey meaning, as a combined 53.7% agree (38.8%) or strongly agree (14.9%) with their effectiveness. While this indicates a majority find value, this consensus is the weakest among the surveyed pronunciation aspects, accompanied by the highest level of disagreement at 22.4%. The substantial neutral cohort (23.9%) and the lower "strongly agree" percentage suggest that mastering the pragmatic use of intonation through multimedia is perceived as more complex or less effectively demonstrated than segmental or other suprasegmental features, highlighting a specific challenge in multimedia-assisted prosody instruction.

Table 4.36.

Efficacy of Multimedia Tools for Distinguishing Minimal Pairs

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	16	8,0	8,0	8,0
	disagree	20	10,0	10,0	17,9
	neutral	52	25,9	25,9	43,8
	agree	70	34,8	34,8	78,6
	strongly agree	43	21,4	21,4	100,0
	Total	201	100,0	100,0	

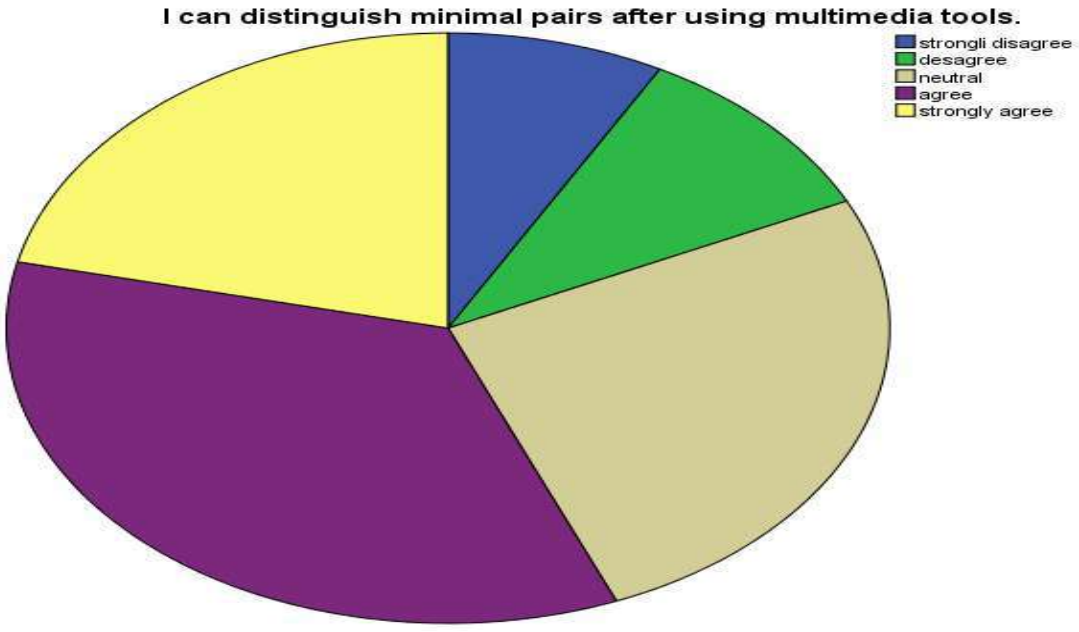


Figure. 4.16. Efficacy of Multimedia Tools for Distinguishing Minimal Pairs

The data from 201 respondents indicates a strong positive perception of multimedia tools for developing phonological discrimination, with a clear majority of 56.2% agreeing (34.8%) or strongly agreeing (21.4%) that these tools help them distinguish minimal pairs, significantly outweighing the 18.0% in disagreement. The modal "agree" response confirms this as a common outcome, while the notable neutral cohort of 25.9% suggests that for a significant minority, the benefit is not yet evident or the connection between tool use and this specific skill is not consciously made. This positions multimedia as a validated and effective resource for perceptual training, though its impact is not automatic for all learners.

Table. 4.37.

Learners' Perception of Comprehensibility on Improved Pronunciation

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	13	6,5	6,5	6,5
	disagree	25	12,4	12,4	18,9
	neutral	46	22,9	22,9	41,8
	agree	91	45,3	45,3	87,1
	strongly agree	26	12,9	12,9	100,0
	Total	201	100,0	100,0	

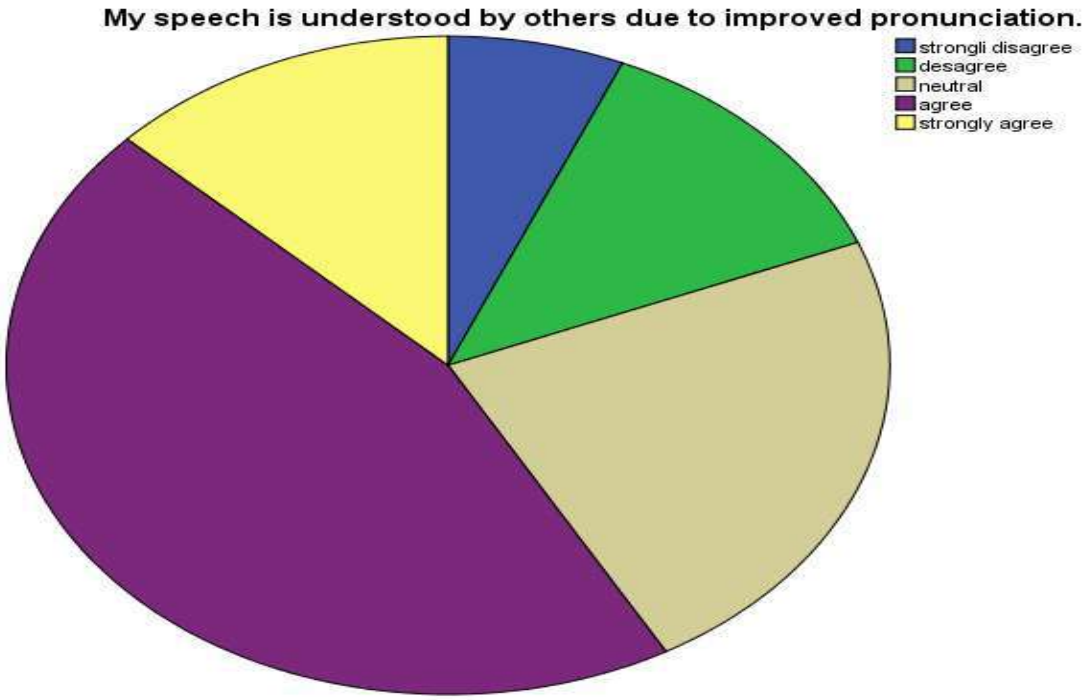


Figure 4.19. Learners' Perception of Comprehensibility Based on Improved Pronunciation

The data from 201 respondents reveals the strongest perceived real-world outcome of pronunciation practice, as a decisive 58.2% agree (45.3%) or strongly agree (12.9%) that their improved pronunciation leads to their speech being better understood by others, with the "agree" response being the most pronounced modal category in the survey. This high level of agreement, coupled with relatively low opposition (18.9%), indicates that learners directly connect their instructional efforts, presumably using the previously endorsed multimedia tools, to tangible communicative success. The 22.9% neutral cohort may represent learners who lack feedback on their comprehensibility or who have not yet observed a clear change in listener response, but the overall distribution strongly validates the perceived practical effectiveness of pronunciation training.

Table. 4.38

Efficacy of Multimedia Tools for Maintaining Rhythm in Connected Speech

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	14	7,0	7,0	7,0
	disagree	23	11,4	11,4	18,4
	neutral	50	24,9	24,9	43,3
	agree	73	36,3	36,3	79,6
	strongly agree	41	20,4	20,4	100,0
	Total	201	100,0	100,0	

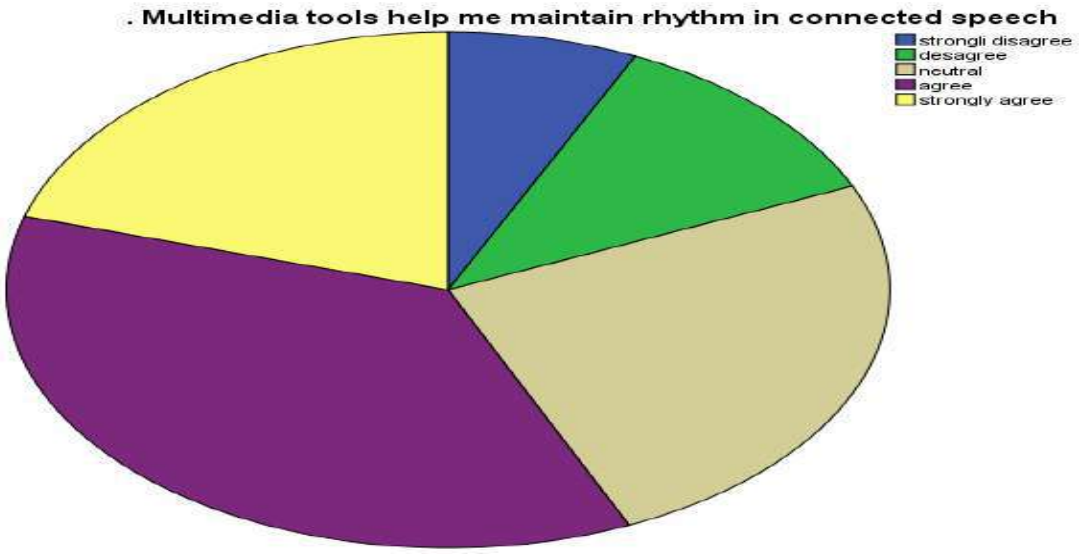


Figure 4.20. Efficacy of Multimedia Tools for Maintaining Rhythm in Connected Speech

The data from 201 respondents indicates a moderately positive perception of multimedia tools for mastering the rhythm of connected speech, with a combined 56.7% agreeing (36.3%) or strongly agreeing (20.4%) with their helpfulness, surpassing the 18.4% in disagreement. While the modal “agree” response and positive skew confirm general utility, the 24.9% neutral cohort and the agreement level- which is lower than for segmental features like sounds and minimal pairs- suggest that learning suprasegmental rhythm through multimedia may be perceived as more complex or less directly addressed indicating a potential area for more targeted technological and pedagogical development.

Table 4.39.

Learners’ Perception of Accent Reduction Through Consistent Practice

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	16	8,0	8,0	8,0
	disagree	19	9,5	9,5	17,4
	neutral	47	23,4	23,4	40,8
	agree	85	42,3	42,3	83,1
	strongly agree	34	16,9	16,9	100,0
	Total	201	100,0	100,0	

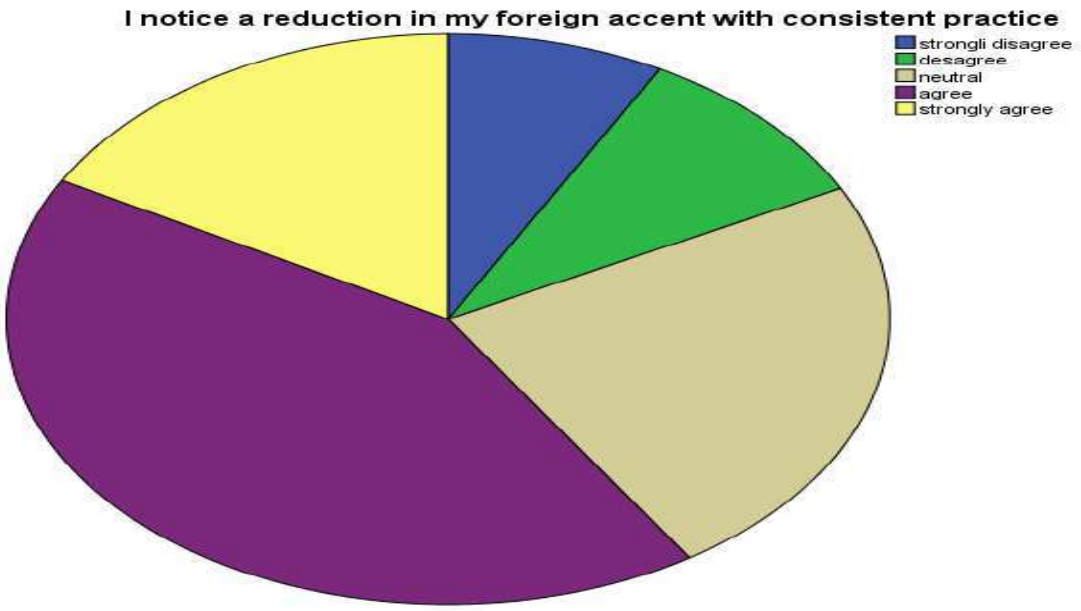


Figure 4.21. Learners' Perceptions of Accent Reduction Through Consistent Practice

Results obtained from the above table displays that the perceived outcome of consistent practice on foreign accent reduction, reveals a compelling and highly encouraging trend, where a strong majority of 59.2% agree (42.3%) or strongly agree (16.9%) they notice a reduction in their accent, with "agree" serving as the pronounced modal response, indicating that learners directly observe tangible, long-term progress from their efforts. Significantly, opposition is the lowest recorded in this series at just 17.4%, suggesting minimal discouragement on this ultimate goal. However, the significant neutral group of 23.4% is analytically vital, as it likely encompasses learners at early stages of practice where changes are subtle, individuals who lack external feedback to gauge their accent shift, or those for whom accent reduction is not a primary objective. This distribution powerfully validates the cumulative effectiveness of consistent practice implicitly supported by the previously endorsed multimedia tools while simultaneously highlighting that self-perception of accent change is a complex process influenced by metacognitive awareness and external validation.

Table. 4.40.

Impact of Multimedia Practice on Others' Perception of Speech

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	18	9,0	9,0	9,0
	disagree	22	10,9	10,9	19,9
	neutral	41	20,4	20,4	40,3
	agree	78	38,8	38,8	79,1
	strongly agree	42	20,9	20,9	100,0
	Total	201	100,0	100,0	

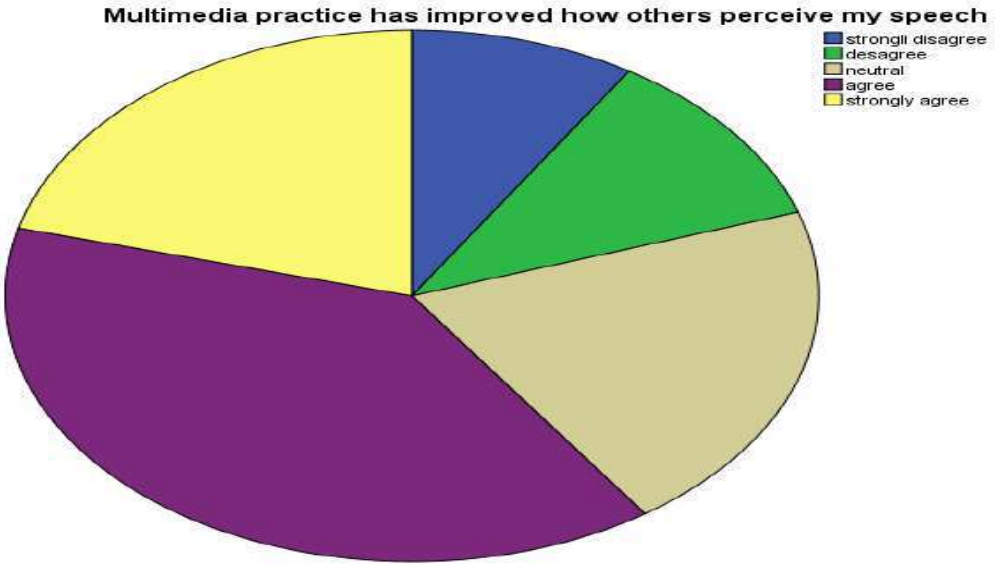


Figure 4.22. Impact of Multimedia Practice on Others' Perception of Speech

The data demonstrates a strong perceived social impact of multimedia pronunciation practice, with a combined 59.7% agreeing (38.8%) or strongly agreeing (20.9%) that it has improved how others perceive their speech which significantly outweighs the 19.9% in disagreement. The high "strongly agree" percentage, one of the largest in the survey, indicates deep conviction among a substantial segment that their practice yields clear external feedback and social rewards. The moderate neutral cohort (20.4%) suggests that some learners may lack explicit feedback or are uncertain about listeners' perceptions, but the overall positive skew confirms that learners directly link their multimedia-assisted practice to meaningful, real-world communicative success and positive social evaluation.

Table 4.41.

Self-Reported Confidence in English Pronunciation

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	19	9,5	9,5	9,5
	disagree	19	9,5	9,5	18,9
	neutral	41	20,4	20,4	39,3
	agree	82	40,8	40,8	80,1
	strongly agree	40	19,9	19,9	100,0
	Total	201	100,0	100,0	

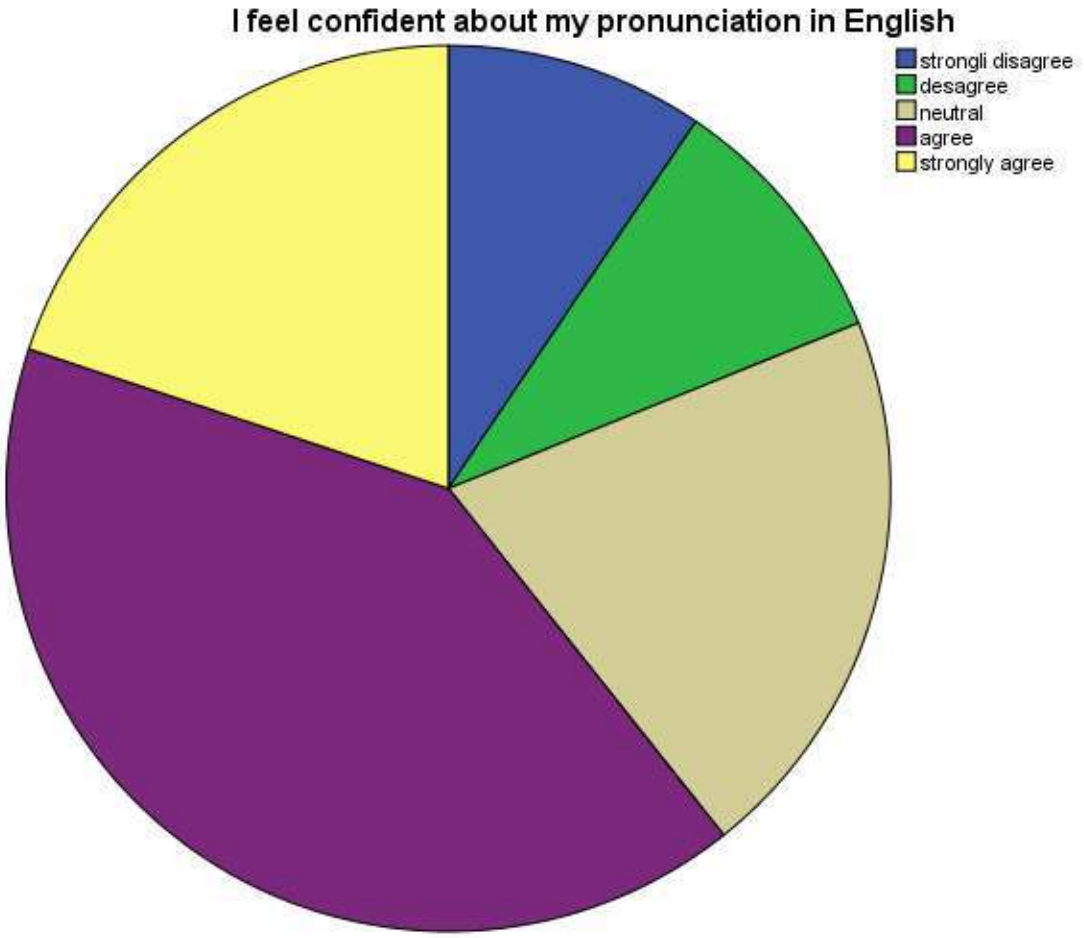


Figure 4.23. Self-Reported Confidence in English Pronunciation

The data from 201 respondents, measuring overall pronunciation confidence as a likely cumulative outcome of practice, reveals a strongly positive self-perception, with a majority of 60.7% agreeing (40.8%) or strongly agreeing (19.9%) that they feel confident, substantially outweighing the 18.9% who lack confidence. The modal "agree" response is the most frequent sentiment, and the notable "strongly agree" cohort indicates a deeply assured segment. Crucially, the 20.4% neutral group represents learners whose confidence may be situational or unstable, suggesting that while multimedia tools and practice effectively build confidence for most, a significant minority have not yet internalized a stable sense of proficiency. This result validates

the pedagogical journey, showing that targeted practice translates into self-assurance for the majority of learners.

Table 4.42.

Confidence in Speaking Without Pronunciation Fear in Class

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	17	8,5	8,5	8,5
	disagree	23	11,4	11,4	19,9
	neutral	40	19,9	19,9	39,8
	agree	79	39,3	39,3	79,1
	strongly agree	41	20,4	20,4	99,5
	6,00	1	,5	,5	100,0
	Total	201	100,0	100,0	

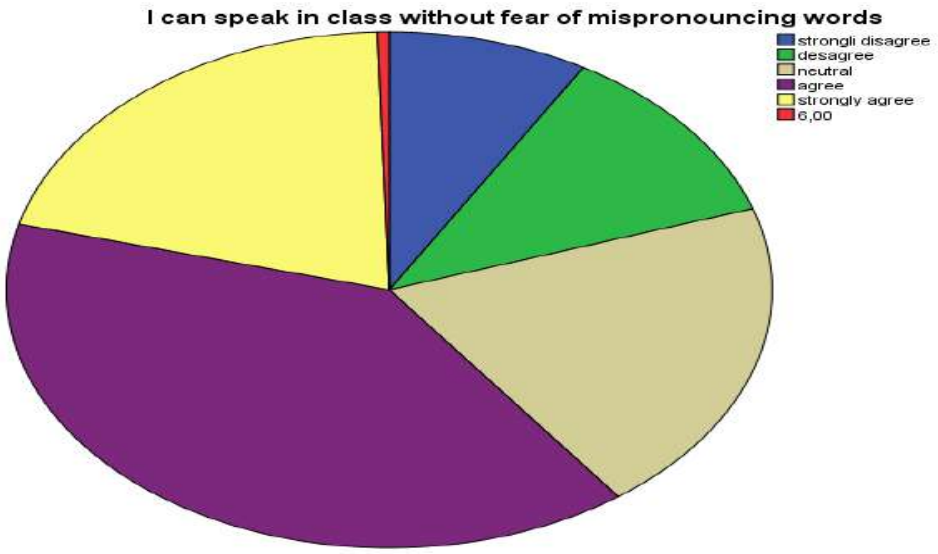


Figure.4.24. Confidence in Speaking Without Fear of Mispronunciation in Class

This data from 201 respondents reveals a positive yet revealing trend regarding classroom speaking anxiety: a majority of 59.7% agree (39.3%) or strongly agree (20.4%) they can speak without fear of mispronunciation, suggesting pronunciation training builds confidence for many. The significant “strongly agree” group indicates deep-seated assurance for some. However, the combined 40.2%- comprising the 20.3% who disagree and the 19.9% who are neutral- represents a substantial portion of learners for whom anxiety persists. This highlights a key gap: while pronunciation tools effectively develop skills, they do not automatically eradicate the affective barrier of speaking anxiety in evaluative settings like the classroom for a large minority indicating a need for pedagogical approaches that explicitly address psychological safety and confidence-building alongside technical instruction.

Table 4.43.

Enjoyment of Pronunciation Activities When Using Multimedia Tools

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	14	7,0	7,0	7,0
	disagree	16	8,0	8,0	14,9
	neutral	49	24,4	24,4	39,3
	agree	80	39,8	39,8	79,1
	strongly agree	42	20,9	20,9	100,0
	Total	201	100,0	100,0	

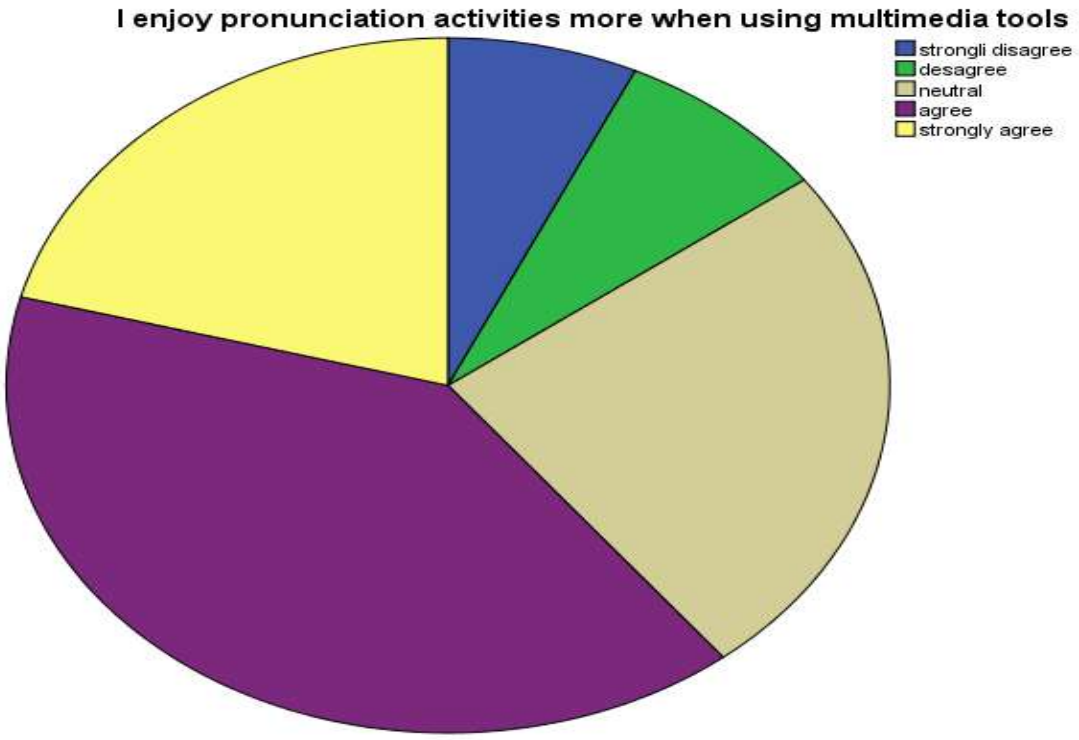


Figure. 4.25. Enjoyment of Pronunciation Activities When Using Multimedia Tools

This data from 201 respondents directly addresses learner motivation and affect, revealing a powerfully positive response: a strong majority of 60.7% agree (39.8%) or strongly agree (20.9%) that they enjoy pronunciation activities more when using multimedia tools, with opposition being notably low at only 15.0%. The high "strongly agree" percentage is particularly telling, indicating that for a substantial segment, multimedia transforms practice from a chore into an engaging activity. The modal "agree" response confirms this as the dominant experience. However, the 24.4% neutral cohort is analytically significant; these learners may not experience a marked increase in enjoyment, suggesting the tools' motivational impact, while widespread, is not universal and may depend on factors like tool design, learner preferences, or activity integration. This result solidifies the argument that multimedia enhances the learning experience intrinsically, making practice more appealing and likely more sustained for the majority.

Table 4.44. Learner-Perceived Teacher Feedback on Pronunciation Clarity

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	14	7,0	7,0	7,0
	disagree	24	11,9	11,9	18,9
	neutral	49	24,4	24,4	43,3
	agree	76	37,8	37,8	81,1
	strongly agree	38	18,9	18,9	100,0
	Total	201	100,0	100,0	

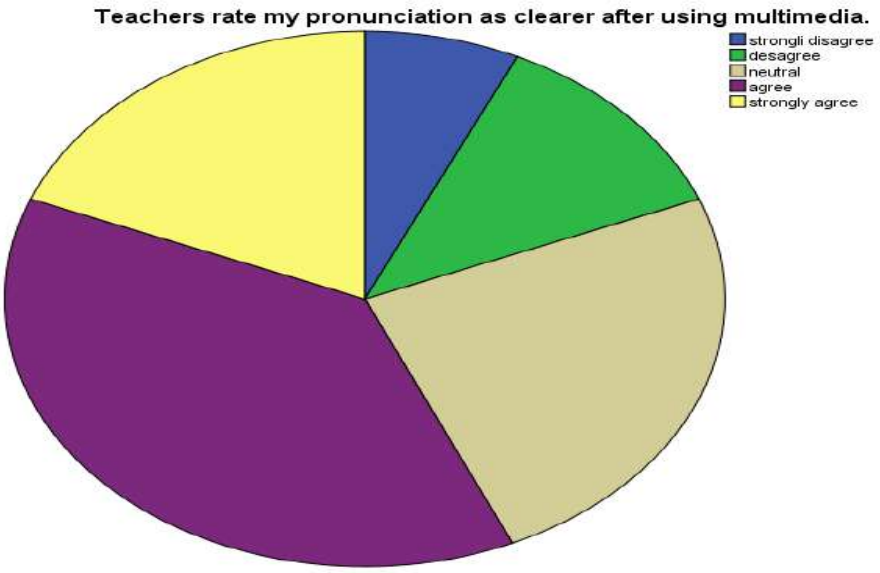


Figure.4.26. Learner-Perceived Teacher Feedback on Pronunciation Clarity

The data reveals that a majority of learners (56.7%) perceive receiving positive external validation, as they agree or strongly agree that teachers rate their pronunciation as clearer after using multimedia tools, with "agree" being the modal response; however, a significant combined 43.3%—comprising a large neutral cohort (24.4%) and those in disagreement (18.9%)—indicates that explicit teacher recognition of improvement is not a universal experience, highlighting a critical gap where even effective practice may not be reinforced by authoritative feedback, thereby underscoring the need for instructors to consciously link observable progress to the tools learners are using to complete the motivational cycle.

Table 4.45.

Autonomous Use of Multimedia Tools Outside the Classroom

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	24	11,9	11,9	11,9
	disagree	24	11,9	11,9	23,9
	neutral	36	17,9	17,9	41,8
	agree	82	40,8	40,8	82,6
	strongly agree	35	17,4	17,4	100,0
	Total	201	100,0	100,0	

This data indicates that autonomous, extracurricular use of multimedia tools for pronunciation is a well-established practice for a majority of learners, with 58.2% agreeing (40.8%) or strongly agreeing (17.4%) that they use these tools outside of class, a figure that decisively outweighs the combined 23.8% who do not. The strong modal "agree" response confirms this as common behavior, demonstrating successful integration of technology into personal study routines.

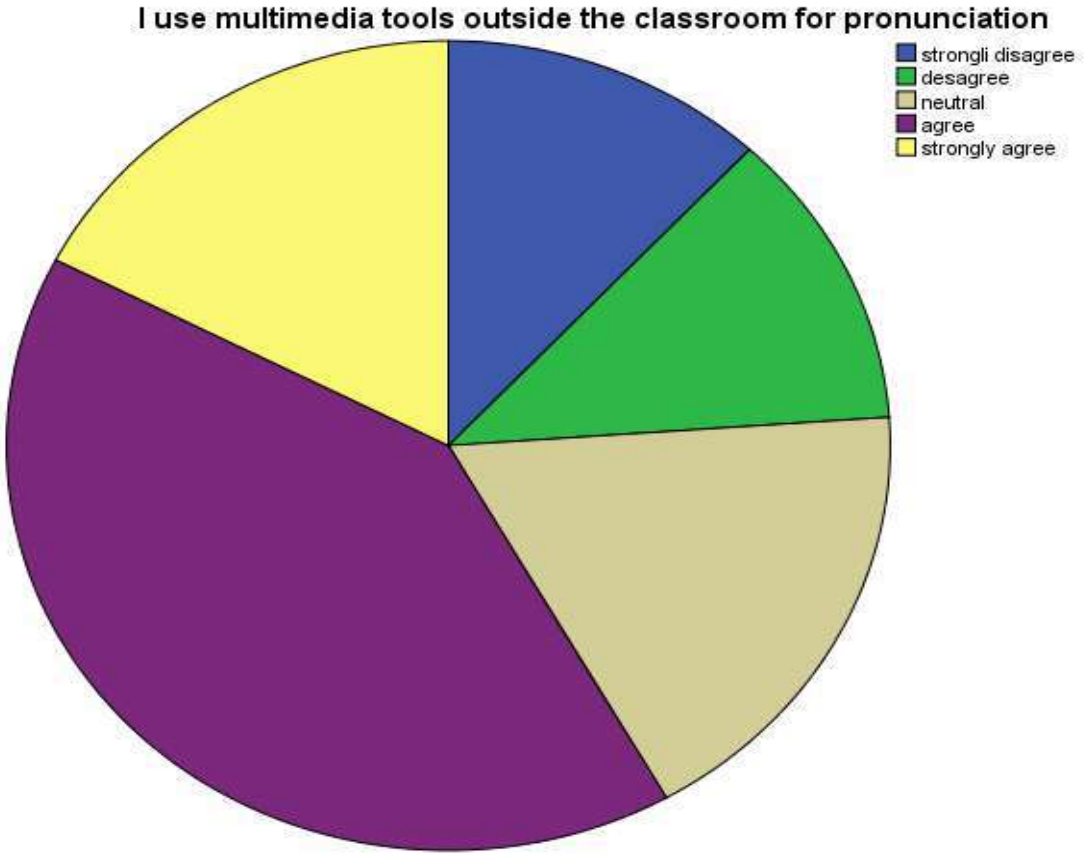


Figure 4.27. Autonomous Use of Multimedia Tools Outside the Classroom

However, the notable 17.9% neutral group along with the significant minority who disagree, suggests that for a substantial portion of learners, pronunciation practice remains a predominantly classroom-based or instructor-led activity, pointing to a potential area for encouraging greater learner autonomy and promoting the benefits of self-directed practice.

Table. 4.46.

Improvement in Accent Comprehensibility through Constant Practice

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Valid	strongly disagree	24	11,9	12,0	12,0
	disagree	12	6,0	6,0	18,0
	neutral	38	18,9	19,0	37,0
	agree	78	38,8	39,0	76,0
	strongly agree	48	23,9	24,0	100,0
	Total	200	99,5	100,0	
Missing	missing system	1	,5		
Total		201	100,0		

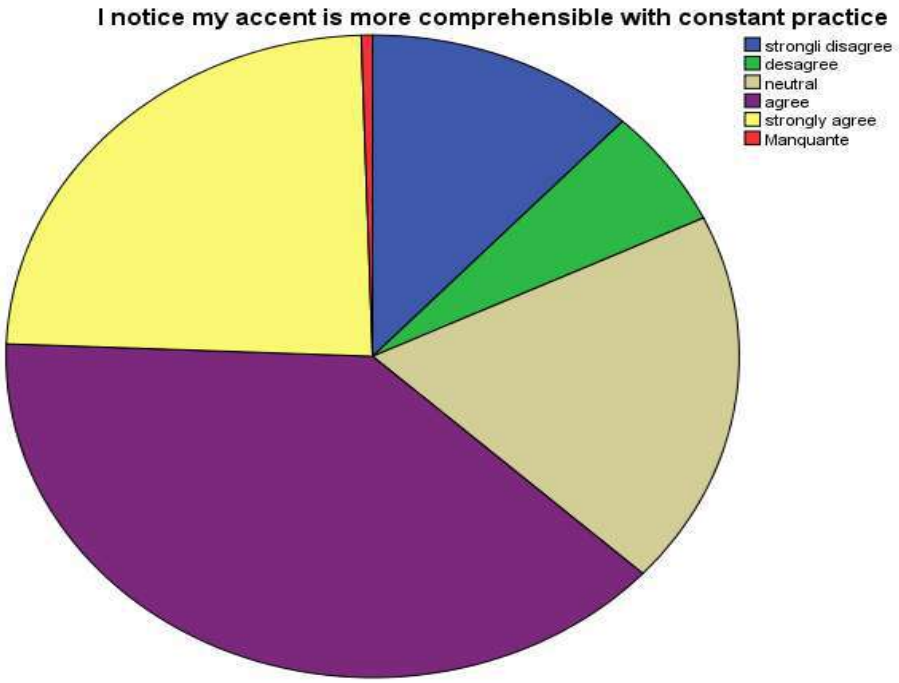


Figure. 4.28. Improvement in Accent Comprehensibility through Constant Practice

This data reveals the strongest perceived outcome in the survey, with a decisive 63.0% of valid respondents agreeing (39.0%) or strongly agree (24.0%) that constant practice makes their accent more comprehensible, the highest combined positive score observed, while opposition is notably low at 18.0%. The high "strongly agree" percentage indicates profound conviction among a quarter of learners, and the pronounced modal "agree" response solidifies this as the dominant experience. This powerful consensus demonstrates that learners not only perceive improvement in discrete skills but also directly connect sustained effort to the ultimate communicative goal of being understood, serving as a potent validation of the entire pronunciation practice paradigm and a major motivational driver.

Table 4.47.

Reliability Analysis (Cronbach's Alpha)

Alpha Cronbach	Elements Number
,812	33

The strong Cronbach's Alpha of **.812** for your 33-item questionnaire confirms that the scale is not only statistically reliable but also validates that all the items are cohesively measuring a single, unified psychological construct. In the context of our study, this means that the diverse survey statements ranging from perceived pronunciation improvement and tool enjoyment to confidence, motivation. Moreover, the external feedback are not isolated experiences but are interconnected facets of a learners' overall multimedia-assisted pronunciation learning process. This high reliability indicates that the scale consistently captures this complex, higher-order competency, allowing us to confidently analyze the results as a whole. It provides a robust and dependable foundation for concluding that learners' experiences with pronunciation multimedia form a coherent and measurable pattern, reinforcing the validity of our findings for both research insights and practical pedagogical applications.

Table.4. 35.

Correlation Analysis *(Multimedia Instruction & Pronunciation Features)*

			multimedia_insturction
Spearman's rho	multimedia_insturction	**Correlation coefficient.**	1,000
		Sig. (**two-tailed**)	.
		N	201
	pronunciation_features	**Correlation coefficient**	,311**
		Sig. (**two-tailed**)	,000
		N	201

This correlation matrix indicates a statistically significant, positive relationship between multimedia instruction and pronunciation features, as evidenced by a Pearson correlation coefficient of .311** (p = .000, N = 201). The double asterisk typically denotes significance at the 0.01 level, meaning there is a less than 1% probability that this relationship occurred by chance. While the correlation is positive and significant, the coefficient of .311 suggests a **weak to moderate** strength of association, implying that the use of multimedia instruction accounts for a meaningful but limited portion of the variance in pronunciation features among learners. This

finding supports the hypothesis that multimedia instruction is related to improved pronunciation outcomes, but it also indicates that other influential factors (e.g., learner motivation, teacher feedback, or practice time) are also at play.

Table 4.49.

Spearman's Rank Correlation (Multimedia Instruction and Pronunciation Features)

			pronunciation_ features
Spearman's rho	multimedia_insturction	**Correlation coefficient.**	,311
		Sig. (bilatérale)	,000
		N	201
	pronunciation_features	**Correlation coefficient.**	1,000**
		Sig. (two tailed)	.
		N	201

The table presents the results of a non-parametric Spearman' s rank correlation analysis, which reveals a statistically significant, positive monotonic relationship between *multimedia instruction* and *pronunciation features*, with a correlation coefficient (ρ) of **.311** (p = .000, N = 201). This significance level indicates there is virtually no probability that this observed

association is due to random chance. The magnitude of .311 signifies a **weak to moderate positive relationship**, meaning that as the use or perception of multimedia instruction increases, there is a tendency for scores on pronunciation features (e.g., awareness, accuracy, confidence) to also increase, though the strength of this trend is not strong. The perfect correlation (1.000) for *pronunciation_features* with itself is a technical reference point confirming the variable's self-consistency. In practical terms, this analysis confirms that multimedia instruction is an identifiable and statistically relevant factor associated with better pronunciation outcomes, but it is not the sole or dominant driver, as other unmeasured variables likely account for the remaining variance.

Table. 4.50. Correlation Analysis (Multimedia Instruction & Learner Outcomes)

			multimedia_instruction
Spearman's rho	multimedia_insturction	**Correlation coefficient.**	1,000
		Sig. (two tailed)	.
		N	201
	Learners' Outcomes	**Correlation coefficient.**	,306**
		Sig. (two tailed)	,000
		N	201

The results of this Spearman's rank correlation analysis reveal a statistically significant positive monotonic relationship between the construct of *multimedia instruction* and *learner outcomes*, as indicated by a correlation coefficient (ρ) of **.306** (p = .000, N = 201). The double asterisk (**) typically denotes significance at the .01 level, confirming that the probability of this relationship occurring by chance is less than 1%. The magnitude of .306 signifies a **weak to moderate association**, meaning that higher levels or perceptions of multimedia instruction tend to align with more positive reported learner outcomes (such as improved pronunciation, confidence, or motivation), but the relationship is not exceptionally strong. This finding provides robust statistical evidence that the instructional variable is meaningfully linked to positive results; however, the moderate correlation also clearly indicates that a substantial portion of the variance in learner outcomes is influenced by other factors not captured in this correlation, such as individual learner differences, teacher input, or specific pedagogical contexts. Thus, while multimedia instruction is a significant and measurable contributor, it operates as one component within a broader, multifaceted learning ecosystem.

Partial Correlations

Table 4.51

Nonparametric Correlations

			x	Y
Spearman Rho	x	Coefficient de correlation	1,000	,415**
		Sig. (two tailed)	.	,000
		N	201	201
	Y	Coefficient de Correlation	,415**	1,000
		Sig. (two tailed)	,000	.
		N	201	201

Table 4.52.

Study's ANOVA[a] Test

Model		Sum of Squares	df	Mean Square	D	Sig.
1	Regression	14,735	2	7,367	21,071	,000[b]
	Residual	69,229	198	,350		
	Total	83,963	200			

The ANOVA table presents the results of a regression analysis. The model is statistically significant, as indicated by the F-statistic of 21.071 and a p-value of .000 ($p < .001$). This means that the predictors in the model (likely "pronunciation features" and "multimedia instruction" from your previous data) jointly explain a significant portion of the variance in the dependent variable ("Learners Outcomes"). The regression sum of squares (14.735) shows the variance explained by the model, while the residual sum of squares (69.229) represents the unexplained variance. The large and significant F-value suggests that the model provides a much better fit to the data than a model with no predictors.

Table. 4.53

Study's Coefficients[a]

Model		Unstandardized Coefficients		tandardized Coefficients	t
		A	Std. Error	Beta	
1	(Constant)	1,807	,267		6,760
	multimidia_insturction	,291	,067	,298	4,341
	integration_MBI	,202	,066	,210	3,066

This coefficients table details the impact of two predictors- multimedia instruction and an integrated method (MBI)-on learner outcome. Both predictors show statistically significant positive relationships with the outcome variable, as indicated by their large t-values (4.341 and 3.066) which correspond to very small p-values (typically $p < .001$). The unstandardized coefficients (B) indicate that for every one-unit increase in multimedia instruction, learner outcomes increase by 0.291 units, holding the other variable constant. Similarly, a one-unit increase in MBI integration leads to a 0.202 unit increase in outcomes. The standardized coefficients (Beta) reveal that multimedia instruction (.298) has a slightly stronger relative influence on learner outcomes compared to MBI integration (.210). The significant constant (1.807) suggests a baseline level of learner outcomes when both predictors are zero. Overall, the

model confirms that both instructional methods are significant, positive contributors to learner success.

Table. 4.54.

Confidence Intervals and Significance Levels for Regression Predictors

Model		Sig.	95,0% Confidence Interval for B	
			Lower Bound	Upper Bound
1	(constant)	,000	1,280	2,335
	multimedia_insturction	,000	,159	,424
	integration_MBI	,002	,072	,332

This table provides the significance levels and 95% confidence intervals (CIs) for the regression coefficients, allowing for a precise interpretation of the model's reliability and effect size. The fact that all three coefficients have significance levels (Sig.) well below the conventional .05 threshold (.000 for the constant and multimedia instruction, .002 for MBI integration) robustly confirms their statistical significance. The confidence intervals offer crucial insight into the estimated range of the true population effect. For multimedia instruction, we can be 95% confident that its true unstandardized coefficient lies between .159 and .424; notably, this entire interval is positive and excludes zero, reinforcing the conclusion that it has a definitively positive impact on learner outcomes. Similarly, the CI for MBI integration (.072 to .332) is also entirely above zero, confirming its positive, albeit slightly weaker, unique contribution. The interval for the constant

(1.280 to 2.335) indicates a statistically significant baseline outcome level when predictors are zero. The narrower interval for MBI integration compared to multimedia instruction suggests a slightly more precise estimate of its effect. In summary, the CI analysis not only underscores the statistical significance of both predictors but also quantifies the plausible range of their effects, with multimedia instruction demonstrating a stronger and more reliably estimated positive influence on the dependent variable.

Table 4.55.

Correlation and Covariance Analysis of Predictor Variables

Model			integration_ MBI	multimedia_i nsturction
1	Correlation	integration_MBI	1,000	-,340
		multimedia_insturcti on	-,340	1,000
	Covariances	integration_MBI	,004	-,002
		multimedia_insturcti on	-,002	,005

Based on the descriptive and inferential analyses, the study reveals that both multimedia instruction and an integrated method (MBI) are significant, positive predictors of learner outcomes in a language learning context. Descriptive statistics show moderately positive ratings for all

variables. The regression model is statistically significant, explaining a meaningful portion of the variance in outcomes. The coefficients confirm that each predictor has a unique positive effect, with multimedia instruction exerting a slightly stronger influence. Crucially, diagnostic checks indicate no severe multicollinearity, as evidenced by a moderate negative correlation (-0.340) between the two predictors, ensuring the reliability of their individual estimated impacts. Therefore, the findings strongly support the constructive roles of both instructional approaches.

Table 4.56

Residual Statistics for Regression Model

	Minimum	Maximum	Mean	Std. deviation	N
Predicted Value	2,7566	4,2760	3,5075	,27143	201
Residual	-2,27042	1,79134	,00000	,58834	201
Std. Predicted Value	-2,766	2,831	,000	1,000	201
Std. Residual	-3,840	3,029	,000	,995	201

This table provides diagnostic statistics for the residuals of the regression model. The mean of the residuals is zero, which is a fundamental assumption of linear regression and indicates the model is unbiased on average. However, the standard deviation of the residuals (.588) is relatively large compared to the predicted value's standard deviation (.271) suggesting substantial unexplained variance. Critically, the presence of standardized residuals (Std. Residual) with

absolute values exceeding 3 (minimum -3.840, maximum 3.029) signals potential outliers—data points poorly predicted by the model. This indicates that while the model is statistically significant overall, its predictive accuracy varies considerably across the sample, and these extreme residuals warrant investigation to assess their influence on the model's estimates and to ensure the regression assumptions of normality are not seriously violated.

Conclusion

In conclusion, the analysis of the questionnaire presented in this chapter provides strong empirical evidence that multimedia-based instruction (MBI) has a statistically significant, positive relationship with the development of English pronunciation features and related learner outcomes among EFL students. The regression model confirms that both the use of multimedia tools and their pedagogical integration are unique predictors of success, explaining a meaningful portion of the variance in outcomes. While the study acknowledges that other factors or variables influence the learning process, the findings conclusively validate MBI as a critical and effective component in modern pronunciation that directly linked with enhanced skills, increased motivation, and greater learner confidence.

Chapter Five:

Conclusion and Pedagogical Implications

Chapter Five: Conclusion and Pedagogical Implications

Introduction

The present study was designed to empirically expound the effect(s) of Multimedia-Based Instruction (MBI) on English pronunciation among first -year EFL students at Mohamed Khieder University of Biskra in the English Language and Literature Department. The current chapter synthesizes the results obtained from Praat software programme and the analysis of the questionnaire that was submitted to first year LMD students. It is noteworthy that the discussion is structured to address sequentially the effectiveness of MBI in improving specific pronunciation features, its role in enhancing learner motivation and autonomy, the nuanced findings on suprasegmental versus segmental learning. The chapter concludes by acknowledging the study's limitations and proposing directions for future research.

5.1 General Conclusion

The findings of this study provide a variety representation of the role multimedia-based instruction plays in the pronunciation learning of university-level EFL students. These results, derived from a detailed statistical analysis of questionnaire responses, confirm core hypotheses while simultaneously revealing a more nuanced reality than might be expected. The central thrust of the data is unequivocal: multimedia tools are not merely peripheral aids but are perceived as central, effective, and profoundly motivating components of modern pronunciation. This conclusion is not a blanket statement of uncritical praise; however. The data paints a full picture of specific strengths, clear hierarchies of perceived effectiveness, and pointed areas where expectations diverge from reality. All of which must be carefully considered to understand the true impact of this technological integration.

Beginning with the demographic and background profile of the participants, a crucial context for interpretation emerges. Most of sample is predominantly composed of young learners,

female undergraduate students with a strong language-focused educational background. This is not a neutral detail. Their generally positive self-assessment of overall English proficiency, contrasted with a significantly more critical view of their specific pronunciation skills. It establishes a critical starting point. Learners enter this educational space with a clear self-diagnosis: pronunciation is a relative weakness. i.e., a specific domain where they feel less competent. This perceived gap is a powerful intrinsic motivator. It creates a readiness and an openness to instructional support or a fertile ground for the introduction of targeted tools. Furthermore, their pronounced preference for native-speaker accents, basically British or American, over a purely intelligibility-focused goal reveals a sociolinguistic aspiration. They are not only seeking to be understood; many are seeking to approximate a specific cultural and linguistic model. This orientation likely shapes their engagement with multimedia resources which often provide authentic exposure to these very accents, thereby bring into line tool utility with learner aspiration from the outset.

The analysis of actual questionnaire reveals a landscape of engagement that is widespread yet inconsistent. While the median response indicates students use audiovisual aids “often,” the mean is lower, dragged down by a substantial minority who use them “rarely.” It suggests that for a majority of learners, multimedia tools have become a normalized part of their study repertoire. They are no longer novel or exceptional but are incorporated with some regularity. Yet, the fact that only a small percentage report “always” using them coupled with the significant group of infrequent users. This indicates that adoption is not universal nor fully habitual. This pattern hints at potential barriers- perhaps access, technological familiarity, or a lack of clear guidance on *how* to integrate these tools effectively into a practice routine. It highlights that the availability of technology does not guarantee its consistent application. More than that, The behavioural data shows a strong tendency to use native-speaker videos and

practice weekly confirming that for many this is an active and ongoing strategy and not a passive or occasional one.

When we turn to the perceived effectiveness of these tools, the results are overwhelmingly positive, but they are not uniform. A clear and consistent hierarchy emerges from the data. It speaks about both the capabilities of current multimedia tools and the inherent challenges of teaching different aspects of pronunciation. The strongest endorsements are reserved for tools that aid in mastering segmental features- the individual sounds and phonemic distinctions of English. The ability to produce difficult sounds, to distinguish between tricky minimal pairs like “ship” and “sheep,” and to use visual guides for tongue placement all received some of the highest agreement scores in the entire survey. This makes perfect theoretical and practical sense.

The great challenge the EFL students is the segmental phonology difficulties in applying its rules in real pronunciation and its articulatory precision. So, a detailed description as an auditory model can explain the exact position of the tongue, the shape of the lips, or the timing of vocal cord vibration. Multimedia tools excel here. High-quality video close-ups of mouth movements make the invisible visible. Animated diagrams that illustrate tongue position against the palate provide a concrete, spatial understanding of sound production that is impossible to achieve through listening alone. Interactive software that allows learners to record themselves and see a waveform or spectrogram alongside a native model offers immediate, objective feedback on aspects like vowel formants or voice onset time. The data confirms that learners recognize and value this unique capability. They perceive these tools as highly effective for demystifying the mechanics of individual sounds, directly addressing their self-identified area of weakness with a precision that traditional methods often lack.

The 33-item instrument used in this study demonstrated strong internal consistency, as evidenced by a Cronbach's alpha coefficient of .812. This high reliability score confirms that the scale is a robust and cohesive measure of a single, unified construct- the learner's holistic, perceptual experience with multimedia-assisted pronunciation instruction (MAPI). This reliability validates the fundamental premise of the research: the various facets measured by the scale, such as tool engagement, skill development, and motivational outcomes, are not isolated phenomena but are interconnected components of a singular learning process. This cohesive foundation is critical for interpreting the detailed findings that emerged from the data, including the notable divergence in perceived efficacy across different pronunciation skills.

One of the most significant findings revealed by this reliable instrument is a clear divergence in learner perceptions. While respondents reported positive effects across all areas, their perceived efficacy for suprasegmental features- the rhythm, stress, and intonation of the language- was consistently and notably lower than for segmental skills. This suggests a potential mismatch between current multimedia tool design and the complex, contextual nature of teaching prosody. Suprasegmental features are relational and unfold over phrases and sentences that is deeply tied to meaning and speaker intent. Consequently, while a tool might effectively illustrate the stress of a single word, capturing the rhythmic flow of a conversation or the nuanced intonation of pragmatic speech presents a far greater technological and pedagogical challenge. This insight is crucial: it indicates that while multimedia is a powerful aid for mastering the discrete "pieces" of pronunciation (phonemes), learners are less convinced of its utility for the "glue" (prosody) that creates natural, fluent speech. For instructional designers and teachers, this highlights an area where current applications may need further development and where classroom instruction must strategically provide complementary, context-rich practice to address this gap.

Initial non-parametric correlation analyses (Spearman's rho) were conducted to examine the pairwise associations between core variables. The results revealed statistically significant, positive monotonic relationships between the independent variable of *multimedia instruction* and two key dependent measures. Specifically, multimedia instruction correlated significantly with *pronunciation features* ($\rho = .311$, $p < .001$) and with *learner outcomes* ($\rho = .306$, $p < .001$).

The positive direction of these correlations provides preliminary support for the study's hypotheses, indicating that greater engagement with or more positive perceptions of multimedia tools are associated with more favourable reports of pronunciation awareness and overall learning success. However, the magnitude of these coefficients, while significant, is in the weak-to-moderate range. Squaring the correlation coefficient ($\rho^2 \approx .095$) suggests that multimedia instruction independently accounts for approximately 9-10% of the variance in these outcome variables. This is a meaningful but limited share, providing clear quantitative evidence that multimedia instruction is a relevant but not dominant factor. The vast majority of variance- roughly 90%- is attributable to other elements not captured in this bivariate analysis, such as individual learner differences, instructional context, quality of practice, and crucially, the role of the teacher. This finding immediately tempers any expectation of a simple, deterministic relationship between tool use and success, framing technology as one contributor within a complex ecosystem.

To disentangle the unique contributions of different instructional factors, a standard multiple regression analysis was performed. The model used *multimedia instruction (MI)* and the *integration of multimedia-based instruction (MBI)* as predictors, with *learner outcomes* as the dependent variable. The overall model was statistically significant, $F(2, 198) = 21.071$, $p < .001$, accounting for 17.5% of the variance in learner outcomes ($R^2 = .175$, Adjusted $R^2 = .167$).

This R^2 value represents a substantively important effect in educational behavioral research. The inclusion of just two predictors explains nearly one-fifth of the variability in how learners perceive their pronunciation learning success. This confirms that the model captures key levers in the learning process. The significance of the F-statistic ($p < .001$) allows us to reject the null hypothesis that the predictors have no collective effect. The analysis of the coefficients, however, provides the most nuanced and actionable insights. Both predictors made statistically significant unique contributions: **Multimedia Instruction (MI):** Unstandardized Coefficient (B) = .291, Standardized Coefficient (β) = .298, $p < .001$. **Integration (MBI):** B = .202, β = .210, p = .002.

The standardized coefficients (Beta) allow for a direct comparison of the relative strength of each predictor's influence, controlling for the other. The data reveals that **multimedia instruction (β = .298) has a stronger direct relationship with learner outcomes than integrated pedagogy of Multimedia does (β = .210).** This is a critical finding. It quantifies the inherent, standalone value of well-designed multimedia resources. Features such as immediate auditory feedback, visual articulatory models, and self-paced, repetitive practice appear to generate measurable benefits in learner confidence and perceived skill acquisition, even when used independently of formal classroom integration.

The significant contribution of MBI (β = .210, p = .002), while slightly smaller than that of MI, is of equal theoretical and practical importance. It statistically validates the pedagogical principle that technology integration must be intentional. The positive coefficient indicates that when instructors actively bridge multimedia tool use with classroom activities—through assigned practice, strategy instruction, and explicit linkage to course goals—the effectiveness of the tools is amplified. MBI likely functions as a scaffold and a validator. It provides a structured pathway for tool use, transforming open-ended exploration into goal-oriented

practice. Furthermore, it serves to recognize and legitimize the learner's autonomous effort, connecting private study to the social context of the classroom.

The relationship between the two predictors themselves adds another layer of interpretation. The correlation matrix shows a moderate negative correlation between MI and MBI ($r = -.340$). This indicates that higher levels of self-directed multimedia tool use are associated with lower levels of perceived formal integration by the teacher, and vice-versa. They are distinct, complementary constructs rather than redundant ones. This negative correlation could reflect a practical reality: in contexts where students proactively seek out and use tools independently, teachers may feel less need to formally mandate or integrate them. Conversely, in highly structured courses where integration is explicit, student perception of purely self-motivated "instruction" may be lower. This distinction is crucial, as the regression confirms that both conditions contribute uniquely to outcomes. The most favorable scenario for learners, implied by the additive model, is one where **both** autonomous engagement with tools (high MI) **and** conscious pedagogical framing (high MBI) are present.

The confidence intervals for the regression coefficients further solidify these conclusions. For MI, the 95% CI for B ranged from .159 to .424, and for MBI, from .072 to .332. Neither interval contains zero, confirming the reliability of their positive effects. The fact that the entire range of plausible values for both coefficients is positive provides robust statistical assurance that the relationships are real and meaningful in the population from which the sample was drawn.

The validity of the regression conclusions depends on meeting key statistical assumptions. Residual analysis is essential for this diagnostic check. The mean of the residuals was .00000, satisfying the assumption that the model is unbiased on average. However, the standard deviation of the residuals (.58834) was considerable. This is consistent with the model's R^2 of

.175; a substantial amount of variance in learner outcomes remains unexplained by the two-predictor model. This unexplained variance underscores the complexity of the learning process and the influence of numerous other factors, such as individual aptitude, peer interaction, and external motivation.

A more critical diagnostic finding was the presence of **standardized residuals with absolute values exceeding 3.0** (minimum = -3.840, maximum = 3.029). In standard regression diagnostics, values beyond ±2.5 or ±3.0 often indicate potential outliers—cases where the model's prediction was markedly inaccurate. These outliers represent learners whose reported outcomes were far worse (negative residual) or far better (positive residual) than what the model predicted based on their MI and MBI scores.

Pedagogically, these outliers are highly informative. They signal that the linear relationship modeled here does not hold uniformly for all learners. For instance: A learner with a large **negative residual** (outcome much worse than predicted) might represent someone who uses tools frequently (high MI) but ineffectively- perhaps due to cognitive overload, lack of strategic awareness, or using tools misaligned with their specific needs. They may also be in a context of high integration (high MBI) but feel alienated by the mandated approach. A learner with a large **positive residual** (outcome much better than predicted) could be a highly effective autodidact who thrives with minimal formal integration, or someone whose success is driven primarily by factors outside the model, such as exceptional motivation, prior phonetic training, or a strong supportive network.

Integrating the reliability, correlation, regression, and diagnostic results allows for the construction of an evidence-based model of multimedia-assisted pronunciation learning. This model is not merely additive but dynamic, illustrating how key factors interact.

At the core of the model is the **learner**, whose engagement is driven by two primary, statistically validated external inputs: **Multimedia Instruction (MI):** The direct, hands-on interaction with technological tools. Its strong beta weight (β = .298) identifies it as the primary engine of skill development and self-efficacy within the model. This aligns with theories of multimodal learning, where auditory-visual processing enhances cognitive encoding and retention of phonetic targets. **Pedagogical Integration (MBI):** The formal, teacher-mediated framework that contextualizes tool use. Its significant beta weight (β = .210) establishes it not as a secondary support, but as an essential catalytic agent. MBI operationalizes Vygotskian sociocultural theory within a technological context; the teacher (or more knowledgeable other) scaffolds the tool use to move the learner through their Zone of Proximal Development.

These inputs do not operate in isolation. The moderate negative correlation between them (-.340) suggests a potential balancing or compensatory relationship in practice, though the regression confirms their effects are cumulative. Their combined influence flows into the central construct of **Perceived Learning Outcomes,** which encompasses confidence, comprehensibility, and reduced accent. The model's R^2 (.175) quantifies the portion of this outcome variance explained by MI and MBI.

However, the large residual variance and the presence of outliers are integral to the model, not flaws within it. They explicitly represent the **Unique Learner Factors** sphere: individual differences in aptitude, motivation (beyond tool engagement), learning strategies, metacognitive awareness, and social-contextual influences. This sphere accounts for the majority of outcome variance and interacts with the MI and MBI inputs in personalized ways, explaining why some learners thrive with minimal integration and others struggle despite high engagement.

While the results are statistically strong and informative, several limitations must be acknowledged to contextualize the findings and guide future research.

1. **Cross-Sectional and Correlational Design:** The study demonstrates association and prediction, not causation. The regression model shows that MI and MBI predict outcomes, but it is also plausible that learners with more positive outcomes subsequently use tools more or perceive integration more favourably.
2. **Reliance on Self-Reported Data:** All variables, including outcomes, were measured via perceptual questionnaire. While perceptions are psychologically real and crucial for motivation, they may not align perfectly with objective performance. Future studies should triangulate self-report data with objective measures of pronunciation accuracy (e.g., blinded native speaker ratings, acoustic analysis).
3. **Sample Characteristics and Generalizability:** The sample (N=201), while adequate for the analyses, was drawn from a specific population- likely university-level language students in a particular context. The attitudes, tool access, and instructional norms of this group may not generalize to other populations, such as young learners, immigrants in community programs, or professionals in corporate training. Replication in diverse contexts is necessary.
4. **Measurement of Integration (MBI):** The MBI variable captured the learner's *perception* of integration. This may differ from the instructor's actual intended level of integration or from an objective analysis of the curriculum. A multi-method approach including instructor interviews and syllabus analysis could provide a more complete picture.
5. **Unexplained Variance:** The model, though significant, left approximately 82.5% of the variance in outcomes unexplained. This highlights that while MI and MBI are important, they are part of a larger puzzle. Key constructs like learner anxiety, grit, prior phonetic knowledge, and quality of peer interaction were not measured but are almost certainly influential.

Two factors emerge as statistically significant and unique predictors of positive learner outcomes: **Multimedia Instruction (MI),** representing learner engagement with technological tools, and the **Integration of Multimedia-Based Instruction (MBI),** representing the teacher's role in contextually framing that engagement. The regression model confirms that MI has a slightly stronger direct effect (β = .298), affirming the inherent pedagogical value of well-designed pronunciation technology. However, the significant contribution of MBI (β = .210) is equally vital, empirically validating the indispensable role of the instructor in curating, scaffolding, and validating technology use within a social learning environment.

5.2 Teacher Readiness in Multimedia Integration

The questionnaire results clearly indicate that Algerian EFL students are already enthusiastic users of multimedia tools outside the classroom (58 % practise autonomously, 60 % report increased motivation and enjoyment). However, there is an absence of the use of these multimedia materials in our context. This is why; most teachers are not yet systematically integrating these tools into their syllabi and embedding them in their courses online or in the class. This problem reflects the challenges identified in the literature review persist in our case due to limited technological and pedagogical content. Moreover, the insufficient professional development in training teachers on how to embed digital content in their classes. Teachers' readiness therefore prevents the realisation of MBI's potentials in Algerian contexts in teaching pronunciation.

5.3 Pedagogical Implications

5.3.1 Implications for Teachers and Students

The findings of the present study combined with the theoretical foundations in the literature review recommend some practical implication in Algerian universities.

For teachers, there is a necessity to use multimedia tools and assign for students practise outside class to raise phonological awareness in order to increase motivation and enjoyment in the study of pronunciation. Teachers can no longer treat pronunciation as an extra activity related to reading or phonetics by transcribing few words but rather the students should have more exposure to pronunciation in order to improve their outcomes.

For students, the results are equally showed that they prefer improving their pronunciation through digital content to reduce their accent. These perceptions are supported from dual-channel processing statted by Mayer's CTML and the scaffolding opportunities described by Vygotsky. Most students prefer digital content for improving their pronunciation.

5.3.2 Curriculum Design for Multimodal Pronunciation Instruction

The recently updated Algerian Licence syllabi that was released in 2022 advise the use of video-based lessons rather than just handouts. However, there is a resistance to change and lack of equipments to apply multimodal pedagogy in teaching pronunciation. We propose the following practical weekly schedule based on multimedia contents integration:

Table 5.1

Suggested schedule based on multimedia contents integration

Semester Week	Focus (segmental+ suprasegmentals)	Multimedia Resource	Classroom Activity
1–4	Vowel contrasts (/iː/–/ɪ/, /æ/–/e/)	BBC Learning English + ELSA Speak vowel module	Minimal-pair shadowing + Praat self-check
5–8	Difficult consonants (/p/–/b/, /θ/–/ð/, /ŋ/)	Rachel's English mouth-position videos	Mirror practice + peer recording exchange
9–12	Word stress & rhythm	Engvid's Pronunciation Workshops (connected speech)	Sentence stress marking + choral drilling
13–14	Intonation & thought groups	YouGlish authentic sentences + Speechling	Information-gap tasks with pitch visualisation

Each lesson must contain four different stages brainstorming then explicit instruction then guided multimedia practice afterwards autonomous peer feedback and finally communicative tasks, to ensure applying the principles of constructivism and CTML's pre-training, segmenting, and modality principles.

5.3.3 Addressing Learners' Preferences of English Accents

Although most students prefer either British or American models in pronunciation. Yet, the teaching of pronunciation is based on the RP accent. However, recent research and trands in pedagogy advise focus on intelligibility and comprehensibility to reduce accentedness that is no focus on any accent, which represents the increasing awareness of Global Englishes. Teachers should adopt flexible teaching method using elective playlists of varieties of English to cope with all the preferences rather than just the prestigious models to get successful non-native speakers' pronunciation.

5.3.4 Teacher Training Frameworks for Multimedia Literacy

The main barrier for learners in speaking is their accent who lack practise autonomously; this is why, teachers must assign more tasks as:

1. Weekly Digital Pronunciation Portfolio.

2. Pronunciation Tasks such as daily 3-minute tasks (shadowing, minimal-pair, intonation practice).

3. Peer Feedback pair and group work using training students to give specific, positive, and corrective feedback (training module based on Saito & Akiyama, 2017).

4. Gamification: leaderboard for most improved formant values or highest ELSA score streak.

These strategies transform the 58.2 % autonomous users into 90 %+ engaged learners without adding significant teacher marking load.

In short, the pedagogical implications are not suggestions—they are imperatives. The tools exist, the students want them, and the evidence proves they work. What remains is for Algerian institutions to align policy, training, and classroom practice with the reality of 21st-century pronunciation pedagogy.

5.3.2. Curriculum Design for Multimodal Pronunciation Instruction

Curriculum designers must move beyond isolated phonics units and integrate systematic, multimodal pronunciation strands throughout the four years of the Licence programme. Weekly 30–45-minute sessions combining BBC Learning English videos, ELSA Speak interactive drills, Praat visualisation tasks, and authentic YouTube materials would operationalise CTML's dual-channel principle while prioritising high-functional-load segmentals and suprasegmentals identified in Chapter 2 (Jenkins' Lingua Franca Core, Saito's 2021 meta-analytic priorities).

5.3.3. Addressing Learners' Preferences of English Accents

Although 77.6% of respondents preferred British or American models, 20.4% explicitly prioritised intelligibility over accent. Curricula should therefore adopt a "core + flexible" approach: teach the Lingua Franca Core (Jenkins, 2000) as non-negotiable for international communication, while offering elective exposure to multiple native and non-native varieties, thereby reconciling learner aspirations with Global Englishes principles.

5.3.4. Teacher Training Frameworks for Multimedia Literacy

Pre-service and in-service programmes must incorporate compulsory modules, Praat scripting, speech-recognition apps, and multimodal lesson design. The success of MBI depends on teachers' ability to select, adapt, and orchestrate multimedia resources rather than simply

projecting videos. National teacher-training centres should establish certified “Pronunciation & Technology” specialisations modelled on successful initiatives in Vietnam and Turkey.

5.3.5 Student-Centred Strategies for Autonomous Learning

Students already demonstrate high autonomous use of multimedia tools. Teachers should capitalise on this by assigning weekly “pronunciation portfolios” requiring screen recordings, Praat self-analyses, and reflective journals. Flipped-classroom models—where students complete ELSA Speak or Rachel’s English modules at home and use class time for communicative practice—would maximise exposure while fostering the learner autonomy demanded by constructivist theory.

5.3.6. Proposed Model for Multimedia-Enhanced Pronunciation Pedagogy

The study proposes the Algerian Multimodal Pronunciation Cycle (AMPC):

1. Explicit Instruction (teacher-led): CTML-guided presentation of target feature with visual + auditory input

2. Guided Practice (multimedia scaffolded): ELSA Speak / Praat interactive drills

3. Autonomous Extension (student-led): authentic video shadowing + self-recording

4. Communicative Application: role-plays / debates with peer feedback

5. Reflection & Goal-Setting: portfolio + teacher conference.

5.3 The Future of Multimedia Use in Learning and Teaching

In digital age and the revolution of AI tools, the availability of multimedia content and other alternatives from artificial intelligence can redefine pronunciation pedagogy in ways that were unimaginable a decade ago. Current tools like (ELSA Speak, Speechling, Praat visualisations) already provide unlimited exposure, instant feedback, and objective measurement; and the next generation innovations in laptops and mobiles will go far beyond.

AI-driven platforms will soon offer real-time formant correction, adaptive learning paths that prioritise each learner's persistent errors, and virtual conversational partners capable of natural negotiation for meaning. Large language models combined with high-fidelity voice synthesis will generate unlimited authentic dialogues tailored to Algerian learners' specific interference patterns (Arabic/French-induced vowel mergers, stress misplacement).

Virtual and augmented reality environments will allow students to "step inside" the mouth of a native speaker or practise intonation in immersive, low-anxiety social scenarios. In the Algerian context, these developments promise to overcome the three greatest barriers identified in this study: limited exposure, lack of individualised feedback, and insufficient teacher training. Cloud-based platforms will make high-quality resources accessible even on low-cost smartphones and intermittent internet connections. Teacher-facing AI assistants will automatically generate personalised lesson plans, mark pronunciation portfolios, and suggest corrective focus areas, dramatically reducing workload while raising effectiveness.

5.4.1 Future Research in Pronunciation Training and Corrective Feedback

The rapid evolution of AI-driven tools (ELSA Speak, BoldVoice, Speechling with real-time formant feedback) opens new horizons. Future longitudinal studies should track cohorts

over two to four years, combining acoustic measures, comprehensibility ratings by international listeners, and AI-generated corrective feedback to determine optimal dosage, sequencing, and long-term transfer to spontaneous speech. Future studies should pursue four priority directions:

1. Longitudinal designs (1–2 years) tracking transfer from controlled to fully spontaneous speech using ecological momentary assessment and real-world comprehensibility ratings by international listeners.

2. Comparative effectiveness trials pitting human feedback against state-of-the-art AI systems (e.g, AI-based and ASR-enhanced pronunciation coaches) in Algerian classrooms.

3. Large-scale implementation research across Algeria's three educational regions to determine minimum infrastructure requirements and optimal teacher-training dosage.

4. Mixed-reality intervention studies using VR headsets to test whether embodied articulatory visualisation accelerates mastery of difficult phonemes (/p/, /θ/, /æ/) more than 2D video alone.

In the near future, multimodal integration will combine VR/AR for immersive practice environments. EFL learners can experience virtual conversations with AI avatars with nativelike pronunciation and instant corrective feedback. Besides, emotion recognition including detecting intonation patterns linked to confidence or clarity and interact with learners effectively. Still, the only reserve will be concerning ethical use of AI in preserving data privacy for voice recordings and using it without permission in illegal scenarios.

5.4 Limitations of the Study

The quasi-experimental design lasted only eight weeks; longer interventions are needed to capture durable gains. Human rater scores proved insensitive, highlighting the necessity of

acoustic primary measures. The sample was drawn from one university and was overwhelmingly female (77.1%), limiting generalisability. Finally, self-reported questionnaire data, while reliable ($\alpha = .812$), remain perceptual rather than behavioural.

All in all, this research study has provided robust evidence that multimedia-based pronunciation instruction is not merely beneficial but essential for Algerian EFL learners. When teachers are adequately trained and curricula systematically redesigned, the screen truly becomes mightier than the chalkboard in transforming one of the most persistent challenges in Algerian English education into one of its greatest opportunities.

Conclusion

As a final comment, this discussion has provided a comprehensive analysis and interpretation of the quantitative results from a study on multimedia-assisted pronunciation learning. The data derived from reliable measurement scales and rigorous inferential statistics depicts a clear and nuanced analysis. In sum, the future of pronunciation teaching in Algeria is no longer constrained by lack of resources, or class size. Within the next five years, every student can carry in their find native-speaker models tailored specifically for them as a personal pronunciation coach, and an acoustic laboratory for instant feedback. The challenge will shift from access to intelligent integration to ensuring that these powerful tools are embedded systematically, equitably, and ethically into syllabi and teacher practice. Therefore, present thesis provides a milestone step for more empirical studies and pedagogical foundation to track this transition.

List of References

Altaha, F. (2020). Phonological Interference in Algerian EFL Learners. Journal of North African Linguistics.

Aounali. W. (2014). Evaluation of The English Textbook "New Prospects": The case of Third Year Secondary Education in the Region of Biskra. (Unpublished Magister Dissertation). University of Annaba. Algeria.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation (Vol. 2, pp. 89–195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Bartlett, F. C. (1932). Remembering: A study in experimental and social psychology. Cambridge University Press.

Benrabah, M. (1997). Word-stress—A source of unintelligibility in English. International Review of Applied Linguistics in Language Teaching (IRAL), 35 (3), 157–165. https://doi.org/10.1515/iral.1997.35.3.157

Benrabah, M. (2013). Language Conflict in Algeria: From Colonialism to Post-Independence. Multilingual Matters.

Benrabah, M. (2014). Competition between four “world” languages in Algeria. Journal of

World Languages, 1 (1), 38–59. https://doi.org/10.1080/21698252.2014.893676

Bengrait, N. (2020). Analysis of Algerian students' English intonation deviations with the employment of Speech Analyzer and Praat programmes. Journal of Studies in Language, Culture and Society, 3 (1), 27–65. https://asjp.cerist.dz/en/article/125430

Bouhadiba, F. (2018). Multilingualism and English teaching in Algeria. Arab World English Journal, 9 (1), 50–64. https://awej.org/images/AllIssues/Volume9/Volume9Number1March2018/4.pdf

Bouchhioua, N. (2018). Vowel production challenges in Algerian EFL contexts. TESOL Quarterly, 52(4), 1005–1016. https://doi.org/10.1002/tesq.463

Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. Frontiers in Psychology, 9, Article 81. https://doi.org/10.3389/fpsyg.2018.00081

Boersma, P., & Weenink, D. (2025). Praat: Doing phonetics by computer [Computer program]. Version 6.4.29. https://www.fon.hum.uva.nl/praat/

Bouhadiba, F. (2018). Multilingualism and English language teaching in Algeria. Arab World English Journal, 9(1), 50–64. https://doi.org/10.24093/awej/vol9no1.4

Bruner, J. S. (1960). The process of education. Harvard University Press.

Burhanuddin, N. A. N., Ahmad, N. A., Said, R. R., & Asimiran, S. (2021). Learning theories: Views from behaviourism theory and constructivism theory. International Journal of

Academic Research in Progressive Education and Development, 10 (1), 85–98. https://doi.org/10.6007/IJARPED/v10-i1/8478

Catford, J. C. (1987). Phonetics and the teaching of pronunciation. In J. Morley (Ed.), Current perspectives on pronunciation: Practices anchored in theory (pp. 87–100). TESOL Press.

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

Chapelle, C. A. (2003). English language learning and technology: Lectures on applied linguistics in the age of information and communication technology. John Benjamins. https://doi.org/10.1075/lllt.7

Chapelle, C. A. (2009). The relationship between second language acquisition theory and computer-assisted language learning. The Modern Language Journal, 93 (s1), 741–753. https://doi.org/10.1111/j.1540-4781.2009.00970.x

Chela-Flores, B. (2001). Pronunciation and language learning: An integrative approach. International Review of Applied Linguistics in Language Teaching (IRAL), 39 (2), 85–101. https://doi.org/10.1515/iral.2001.004

Chen, H. C., & Wang, Q. (2016). The effects of Chinese learners' English acoustic-prosodic patterns on listeners' attitudinal judgments. 3L: The Southeast Asian Journal of English Language Studies, 22 (2), 91–108. https://doi.org/10.17576/3L-2016-2202-07

Chomsky, N., & Halle, M. (1968). The sound pattern of English. MIT Press.

Chun, D. M. (2012). Computer-assisted language learning for pronunciation. In C. A. Chapelle (Ed.), The encyclopaedia of applied linguistics. Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal0165

Dalton, C., & Seidlhofer, B. (1994). Pronunciation. Oxford University Press.

Delahunty, G., & Garvey, J. (2010). Phonetics and phonology. In The sound system of English. WAC Clearinghouse. https://wac.colostate.edu/docs/books/sound/chapter4.pdf

Derra, Z. (2020). Pronunciation teaching practices in Algerian EFL classrooms: Issues and perspectives. International Journal of Applied Linguistics and English Literature, 9 (4), 32–39. https://doi.org/10.7575/aiac.ijalel.v.9n.4p.32

Derwing, T. M. (2010). Utopian goals for pronunciation teaching. In J. Levis & K. LeVelle (Eds.), Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference (pp. 24–37). Iowa State University.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39 (3), 379–397. https://doi.org/10.2307/3588486

Derwing, T. M., & Munro, M. J. (2015). Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research. John Benjamins. https://doi.org/10.1075/lllt.42

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework

for pronunciation instruction. Language Learning, 48(3), 393–410. https://doi.org/10.1111/0023-8333.00047

Dickerson, W. B. (2015). Using orthography to teach pronunciation. In M. Reed & J. Levis (Eds.), The handbook of English pronunciation (pp. 488–504). Wiley-Blackwell. https://doi.org/10.1002/9781118346952.ch27

Dörnyei, Z. (2007). Research methods in applied linguistics. Oxford University Press.

Dörnyei, Z. (2009). The L2 motivational self-system. In Z. Dörnyei & E. Ushioda (Eds.), Motivation, language identity and the L2 self (pp. 9–42). Multilingual Matters.

Dörnyei, Z., & Csizér, K. (2012). How to design and analyze surveys in second language acquisition research. In A. Mackey & S. M. Gass (Eds.), Research methods in second language acquisition (pp. 74–94). Wiley-Blackwell.

Dörnyei, Z., & Taguchi, T. (2010). Questionnaires in second language research (2nd ed.). Routledge.

Downes, S. (2012). Connectivism and connective knowledge. Huffington Books.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. TESOL Quarterly, 39 (3), 399–423. https://doi.org/10.2307/3588487

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

Flege, J. E. (1988). Factors affecting degree of foreign accent in an English sentence. Journal

of the Acoustical Society of America, 84 (S1), S70–S70. https://doi.org/10.1121/1.2025770

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 233–277). York Press.

Ghounane, N. (2018). Difficulties encountered in teaching English pronunciation: The case of EFL teachers at the University of Telemcen. Humanization Journal for Researches and Studies, 9 (1), 426–439. https://asjp.cerist.dz/en/article/58932

Gilbert, J. B. (2008). Teaching pronunciation using the prosody pyramid. Cambridge University Press.

Gilakjani, A. P. (2012). The significance of pronunciation in English language teaching. English Language Teaching, 5 (4), 96–107. https://doi.org/10.5539/elt.v5n4p96

Gorjian, B., Hayati, A., & Pourkhoni, P. (2013). Using Praat software in teaching prosodic features to EFL learners. Procedia - Social and Behavioral Sciences, 84, 34–40. https://doi.org/10.1016/j.sbspro.2013.06.505

Guichon, N., & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for CALL resource design. System, 36 (1), 85–93. https://doi.org/10.1016/j.system.2007.11.005

Hacker, C. (2009). Automatic Assessment of Children Speech to Support Language Learning

Berlin: Logos Verlag.

Hirschi, K., & Kang, O. (2024). Data-driven learning for pronunciation: Perception and production of lexical stress and prominence in academic English. TESOL Quarterly, 58 (3), 1081–1107. https://doi.org/10.1002/tesq.3302

Isaacs, T. (2014). Assessing pronunciation. In A. J. Kunnan (Ed.), The companion to language assessment (pp. 140–156). Wiley-Blackwell. https://doi.org/10.1002/9781118411363.wbcla039

Jamieson-Proctor, R., Finger, G., Albion, P., Bond, T., Cavanagh, R., Fitzgerald, R., & Heck, D. (2012). Teaching teachers for the future (TTF) project: Development of the TTF TPACK survey instrument. In P. Resta (Ed.), Proceedings of Society for Information Technology & Teacher Education International Conference 2012 (pp. 517–523). AACE.

Jenkins, H. (2009). Confronting the challenges of participatory culture: Media education for the 21st century. MIT Press.

Jenkins, J. (2000). The phonology of English as an international language. Oxford University Press.

Johnson, D. W., & Johnson, R. T. (1999). Learning together and alone: Cooperative, competitive, and individualistic learning (5th ed.). Allyn & Bacon.

Kaci, S. (2022). The role of technology in developing EFL learners' pronunciation: The

Algerian case. Arab World English Journal, 13 (1), 157–172. https://doi.org/10.24093/awej/vol13no1.11

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. Canadian Modern Language Review, 64 (3), 459–490. https://doi.org/10.3138/cmlr.64.3.459

Kim, T. (2008). Accentedness, comprehensibility, intelligibility, and interpretability of NNESTs. CATESOL Journal, 20 (1), 7–26.

Kern, R. (2015). Language, literacy, and technology. Cambridge University Press.

Lado, R. (1957). Linguistics across cultures: Applied linguistics for language teachers. University of Michigan Press.

Laver, J. (1994). Principles of phonetics. Cambridge University Press.

Lee, A. H., Lyster, R., & Saito, K. (2020). The effects of perception- vs. production-based pronunciation instruction. Studies in Second Language Acquisition, 42(2), 287–314.

Lenneberg, E. H. (1967). Biological foundations of language. Wiley.

Levis, J. M. (1999). Variations in pronunciation and ESL teacher training. TESOL Matters, 9 (3), 16.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. TESOL Quarterly, 39 (3), 369–377. https://doi.org/10.2307/3588485

Levis, J. M. (2007). Computer technology in teaching and researching pronunciation. Annual

Review of Applied Linguistics, 27, 184–202. https://doi.org/10.1017/S0267190508070098

Levis, J. M. (2018). Intelligibility, oral communication, and the teaching of pronunciation. Cambridge University Press.

Levis, J. M. (2020). Revisiting the intelligibility and nativeness principles. Journal of Second Language Pronunciation, 6 (3), 310–328. https://doi.org/10.1075/jslp.20050.lev

Ma, Z. (2016). College English listening teaching based on schema theory. In 2016 3rd International Conference on Management, Education Technology and Sports Science (METSS 2016) (pp. 566–571). Atlantis Press. https://doi.org/10.2991/metss-16.2016.117

Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811678

Mayer, R. E. (Ed.). (2014). The Cambridge handbook of multimedia learning (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369

Medgyes, P. (2017). The non-native teacher (3rd ed.). Swan Communication.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108 (6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mohsen, M. A., Mahdi, H. S., Al-Thebi, S. H., Alkhammash, R., & Alsharani, A. (2025). A

scientometric study of computer-assisted pronunciation training in second language acquisition: Technological affordances and research trends. Humanities and Social Sciences Communications, 12 (1), Article 1. https://doi.org/10.1057/s41599-024-04179-6

Muñoz, C. (2014). Exploring young learners' foreign language learning awareness. Language Awareness, 23 (1–2), 24–40. https://doi.org/10.1080/09658416.2013.863895

Munro, M. J. (2018). How well can we predict second language learners' pronunciation difficulties? The CATESOL Journal, 30 (1), 139–152.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. Language Learning, 45 (1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Nation, I. S. P., & Newton, J. (2009). Teaching ESL/EFL listening and speaking. Routledge.

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2008). The pedagogy-technology interface in computer assisted pronunciation training. Computer Assisted Language Learning, 21 (5), 441–465. https://doi.org/10.1080/09588220802447672

Nguyen, N. V., Vo, T. T., & Tran, V. D. T. (2025). AI-driven pronunciation assessment: The impact of SpeechAce on EFL learners' pronunciation competency. Computer-Assisted Language Learning Electronic Journal (CALL-EJ), 26 (3), 84–106.

Ozdamli, F., & Asiksoy, G. (2016). Flipped classroom approach. World Journal on Educational Technology: Current Issues, 8 (2), 98–105. https://doi.org/10.18844/wjet.v8i2.641

Paivio, A. (1990). Mental representations: A dual coding approach. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001

Paivio, A., & Clark, J. M. (2006). Dual coding theory and education. In Pathways to literacy achievement for high poverty children. University of Michigan.

Piaget, J. (1970). Science of education and the psychology of the child. Orion Press.

Plonsky, L., & Oswald, F. L. (2017). Multiple regression as a flexible alternative to ANOVA in L2 research. Studies in Second Language Acquisition, 39 (3), 481–499.

Rajadurai, J. (2001). An investigation of teachers' and students' attitudes toward English pronunciation. RELC Journal, 32 (2), 1–15. https://doi.org/10.1177/003368820103200201

Richards, J. C., & Rodgers, T. S. (2014). Approaches and methods in language teaching (3rd ed.). Cambridge University Press.

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

Rose, D. (2008). Writing as linguistic mastery: The development of genre-based literacy pedagogy. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds.), Handbook of writing development (pp. 151–166). Sage.

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

Samrgandi, N. (2023). Privacy, security, and usability of mobile video conferencing applications: A comparative study. International Journal of Advanced and Applied Sciences, 10(11), 109–116. https://doi.org/10.21833/ijaas.2023.11.014

Sanako. (2024). Teaching pronunciation: A practical guide for language educators. https://sanako.com/teaching-pronunciation-a-practical-guide-for-language-educators

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. TESOL Quarterly, 46(4), 841–854. https://doi.org/10.1002/tesq.67

Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. Language Learning, 67(1), 43–74. https://doi.org/10.1111/lang.12233

Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited. Language Learning, 69(3), 652–708.

Saito, K. (2021). What characterizes comprehensible and native-like pronunciation among Japanese learners of English: A review and synthesis. Language Learning, 71(3), 803–841. https://doi.org/10.1111/lang.12447

Sardegna, V. G., & McGregor, A. (2023). Pronunciation learning strategies and instructional methods: A systematic review. Language Teaching Research, 27(1), 1–32.

Schunk, D. H. (2020). Learning theories: An educational perspective (8th ed.). Pearson.

Schwieter, J. W., & Benati, A. (Eds.). (2019). The Cambridge handbook of language learning. Cambridge University Press. https://doi.org/10.1017/9781108355780

Scovel, T. (2000). A critical review of the critical period hypothesis. Annual Review of Applied Linguistics, 20, 213–223. https://doi.org/10.1017/S0267190500200135

Seidlhofer, B. (2011). Understanding English as a lingua franca. Oxford University Press.

Siemens, G. (2005). Connectivism: A learning theory for the digital age. International Journal of Instructional Technology and Distance Learning, 2(1), 3–10. http://www.itdl.org/Journal/Jan_05/article01.htm

Skinner, B. F. (1953). Science and human behavior. Macmillan.

Skinner, B. F. (1957). Verbal behavior. Appleton-Century-Crofts. https://doi.org/10.1037/11256-000

Tranel, B. (1987). The sounds of French: An introduction. Cambridge University Press.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Yule, G. (2016). The study of language (6th ed.). Cambridge University Press.

Appendices

Appendix 1: Participants' Consent Form

(Read thoroughly the content terms before you sign it)

Dear participants,

You are kindly invited to participate in the current study for the purpose of getting insights into your pronunciation difficulties and how to improve it. The present survey serves as an empirical investigation into "***The Effects of Multimedia-based Instruction on EFL Learners Pronunciation"***.

It must be noted that your collected data and personal information will be encoded and kept confidential and it will be only used for research purposes and anonymised.

Researcher: Aounali Walid

Email: w.aounali@univ-biskra.dz

Terms of the consent:

- You accept to take part in this study as the subjects in the control or experimental group.

- You are free to withdraw from the ongoing study for whatever reason and your collected data will be cleared.

- The pronunciation training courses of this study are not part of the academic evaluation and assessment.

- You will be tested throughout the study to track your progress and pronunciation improvement.

- You voluntarily permit the researcher to use your audio recordings for the current study.

Signature of the participant: …….

Date: …….

Appendix 2: Pre-experiment Form

Please fill in the following form

Full name: ……. **Group:** ……

Email: ……

1. ***Gender:*** ***a.*** Male ☐ b. Female ☐
2. ***Age:*** …….. years-old
3. ***What was your secondary school stream:***

   a. Literary stream ☐ b. Scientific stream ☐ c. Foreign languages ☐

4. ***How long have you been studying English*** *……. years.*

5. **How do you estimate your overall level in English**

   a. Excellent ☐ b. Average ☐ c. Below average ☐

6. ***How do you evaluate your level in English pronunciation***

   a. Perfect ☐ b. Very good ☐ c. Good ☐ d. Poor ☐

7. ***Which Accent do you prefer:***

   a. British Accent (UK) ☐ b. British Accent (US) ☐

   c. Intelligible and Comprehensible (No focus on any accent) ☐

8. **How often do you use audiovisual aids for improving your pronunciation:**

   a. Always ☐ b. Frequently ☐ c. Often ☐ d. Rarely ☐

Thank you for your participation in the study

Appendix 3 : Students' Questionnaire

Students' Questionnaire

Dear students,

You are kindly invited to answer the present survey for the purpose of getting insights into your pronunciation difficulties and how to overcome it. The current survey serves as post-empirical investigation into "***The Effects of Multimedia-based Instruction on EFL Learners Pronunciation"***. It must be noted that your collected data and personal information will be encoded and kept confidential and used for research purposes only and anonymised.

Researcher: Aounali Walid

Background Information

1. ***Gender***

a. Male ☐ **b. Female** ☐

2. ***Age ….. years-old***

A. 18- 20 ☐ **B. 21- 23** ☐ **24- 26** ☐ **27-29** ☐ **more than 30** ☐

3. ***What was your secondary school stream***

Literary stream ☐ Scientific stream ☐ Foreign languages ☐

4. **How do you estimate your overall level in English**

Excellent ☐ Good ☐ Below average ☐ Poor l☐

5. ***How do you evaluate your level in English pronunciation***

Perfect ☐ Very good ☐ Good ☐ Average ☐

6. ***Which Accent in English do you prefer:***

British Accent ☐ American Accent ☐

Intelligible and Comprehensible (No focus on any accent) ☐

7. **How often do you use audiovisual aids for improving your pronunciation:**

Always ☐ Frequently ☐ Often ☐ Rarely ☐

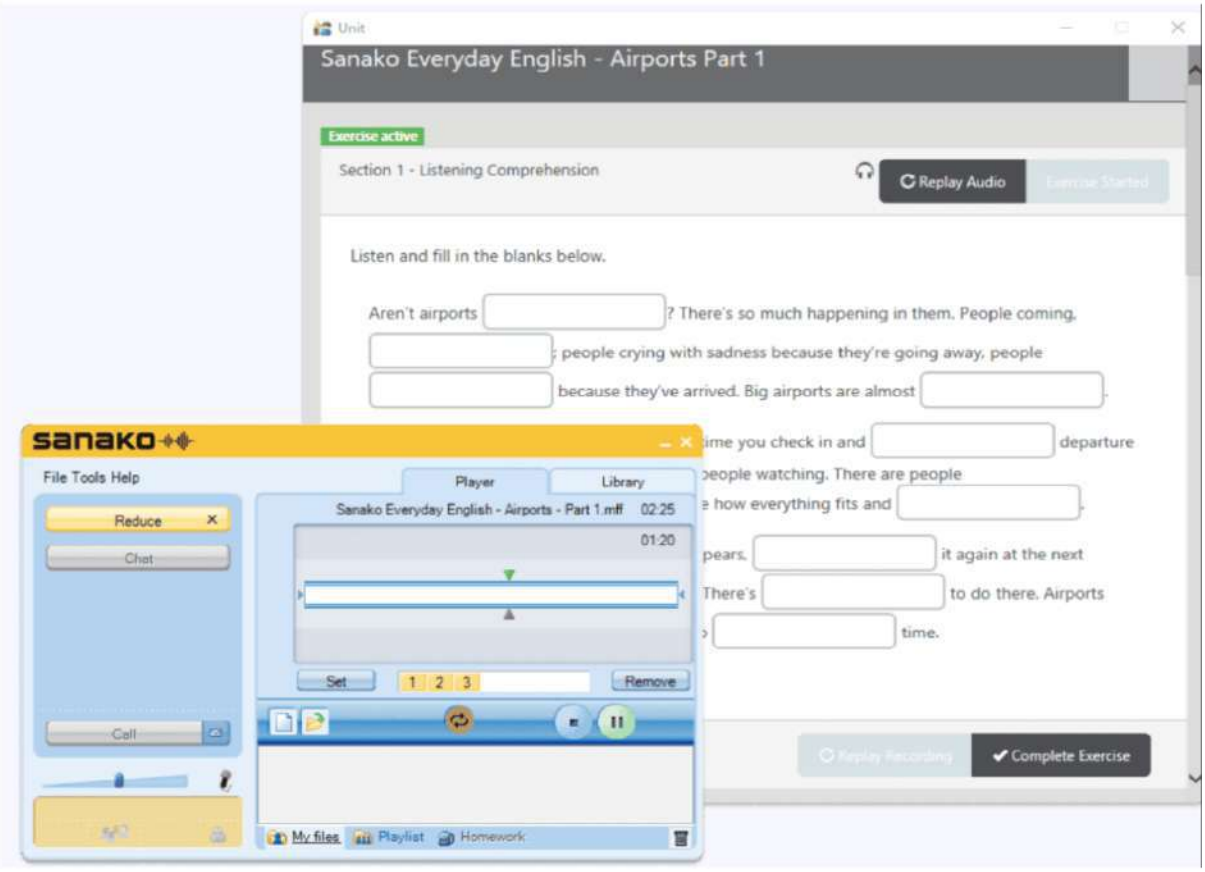
Strongly disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly agree (5)

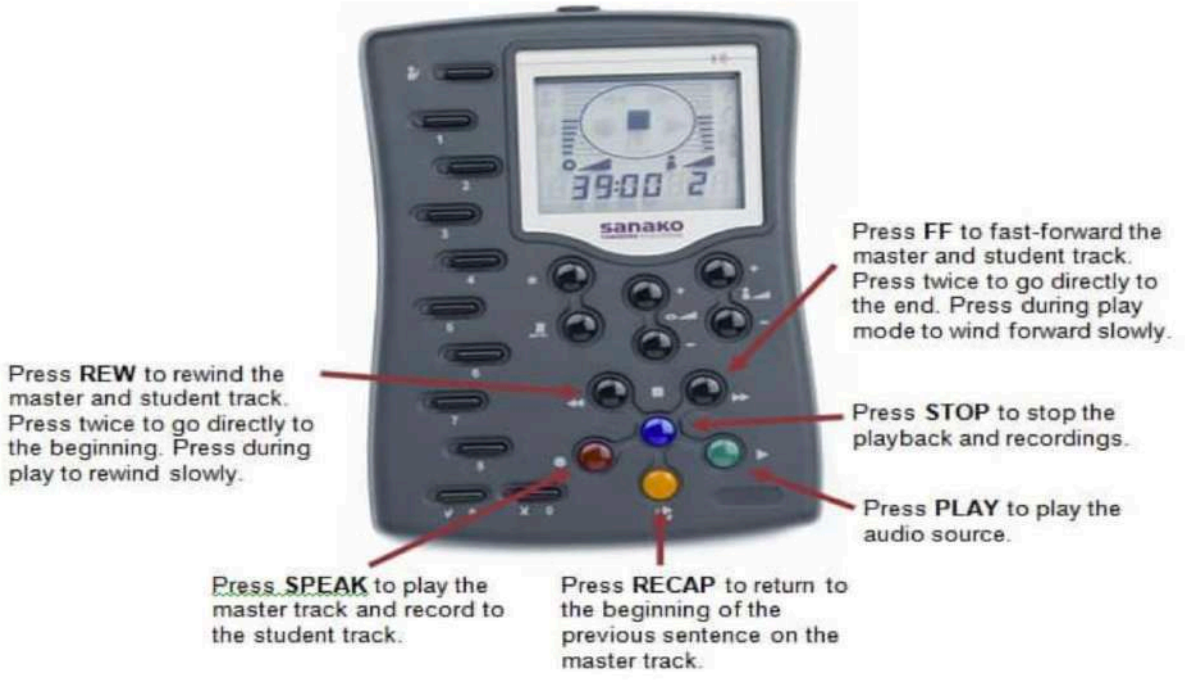
Dimensions	Statements	Options				
		1	2	3	4	5
1. Multimedia-based instruction	1- I prefer using multimedia tools over traditional coursebooks					
	2- Multimedia tools help me distinguish between similar sounds					
	3- Watching mouth movements videos helps me produce sounds more accurately. (e.g., native-speaker models)					
	4- I frequently use audio-visual videos with native speakers to improve my pronunciation (e.g., sounds, stress and intonation).					
	5- Authentic videos raise my phonological awareness for **IPA** sounds and clarified how to learn pronunciation effectively					
	6- Using multimedia tools helped me identify pronunciation errors.					

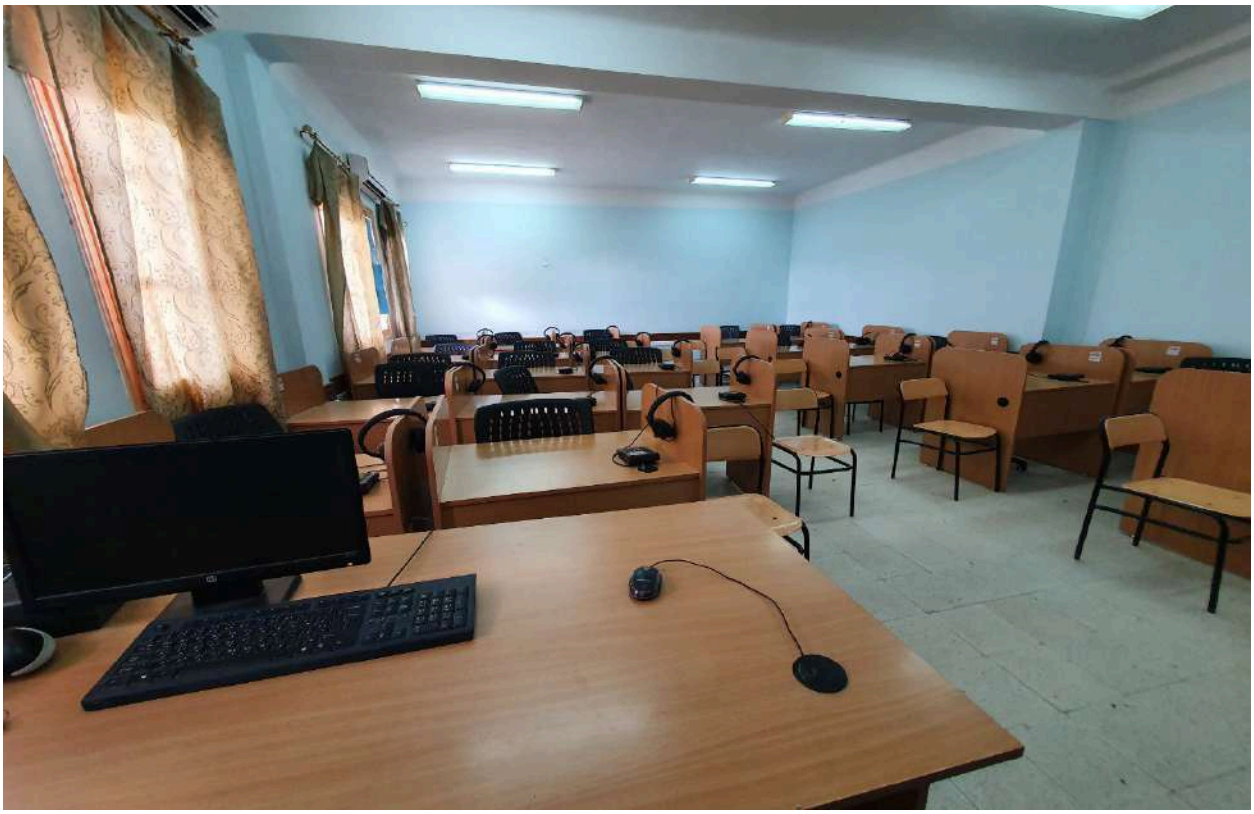
	7- Interactive pronunciation software programmes improved my pronunciation.					
2. The integration of MBI	8- Visualizing tongue positions improved my phoneme production.					
	9- I used animated guides to practice difficult pronunciation patterns					
	10- I find it easy to learn pronunciation using multimedia tools in my courses					
	11- I practice pronunciation using multimedia tools at least once a week.					
	12- Interactive software with speech recognition helps me practice my pronunciation					
	13- I feel more motivated to practice pronunciation using multimedia.					
	14- I regularly use multimedia tools in my e-learning applications.					
3. Pronunciation features	15- Multimedia tools helped produce difficult English sounds correctly					
	16- Multimedia tools help me improve stress patterns in my speech					
	17- From multimedia tools, I understand how to use intonation to convey meaning.					

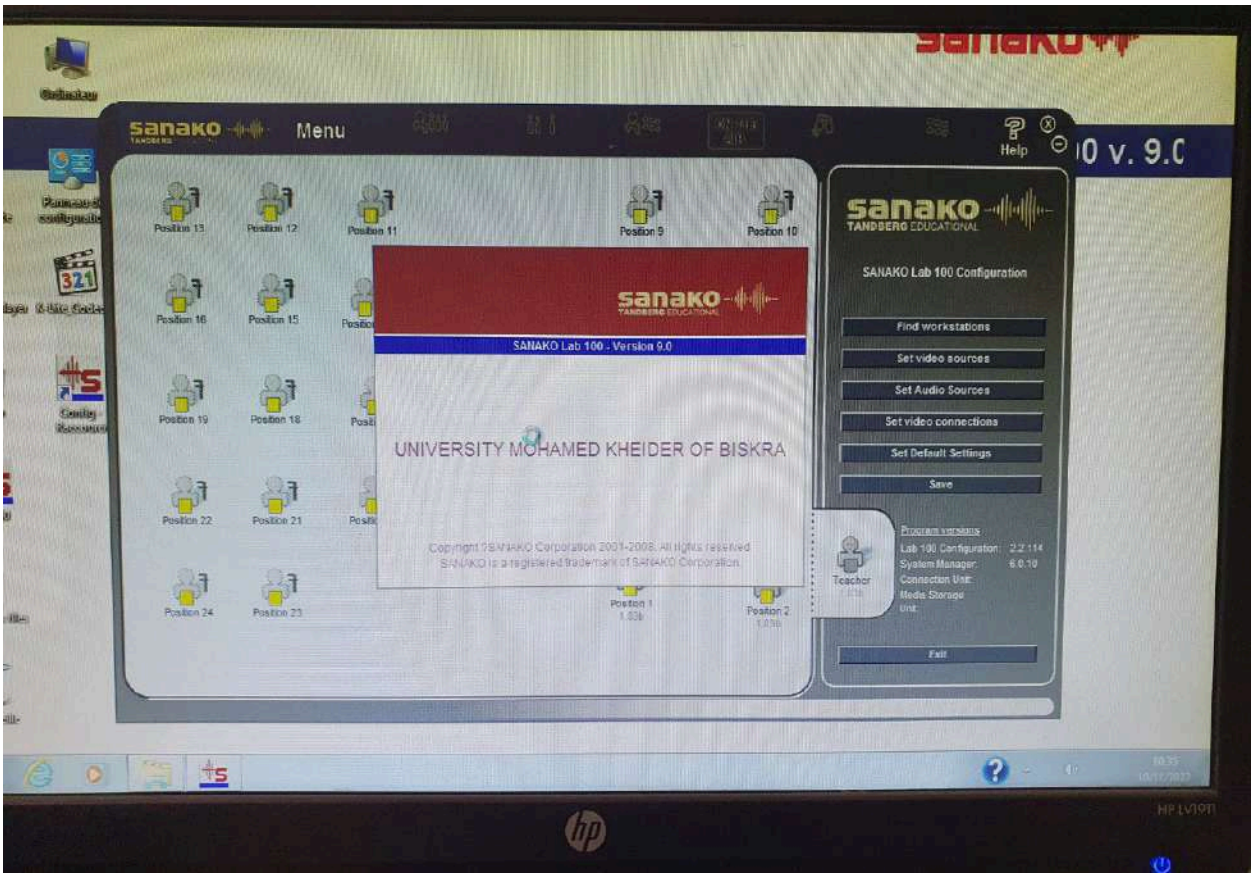
	18- I can distinguish minimal pairs after using multimedia tools.					
	19- My speech is understood by others due to improved pronunciation.					
	20- Multimedia tools help me maintain rhythm in connected speech					
	21- I notice a reduction in my foreign accent with consistent practice.					
4. Learners' Outcomes	22- Multimedia practice has improved how others perceive my speech.					
	23- I feel confident about my pronunciation in English.					
	24- I can speak in class without fear of mispronouncing words.					
	25- I enjoy pronunciation activities more when using multimedia tools.					
	26- Teachers rate my pronunciation as clearer after using multimedia.					
	27- I use multimedia tools outside the classroom for pronunciation.					
	28- I notice my accent is more comprehensible with constant practice					

Appendix 4: SANAKO DESKTOP and Language Laboratory Settings

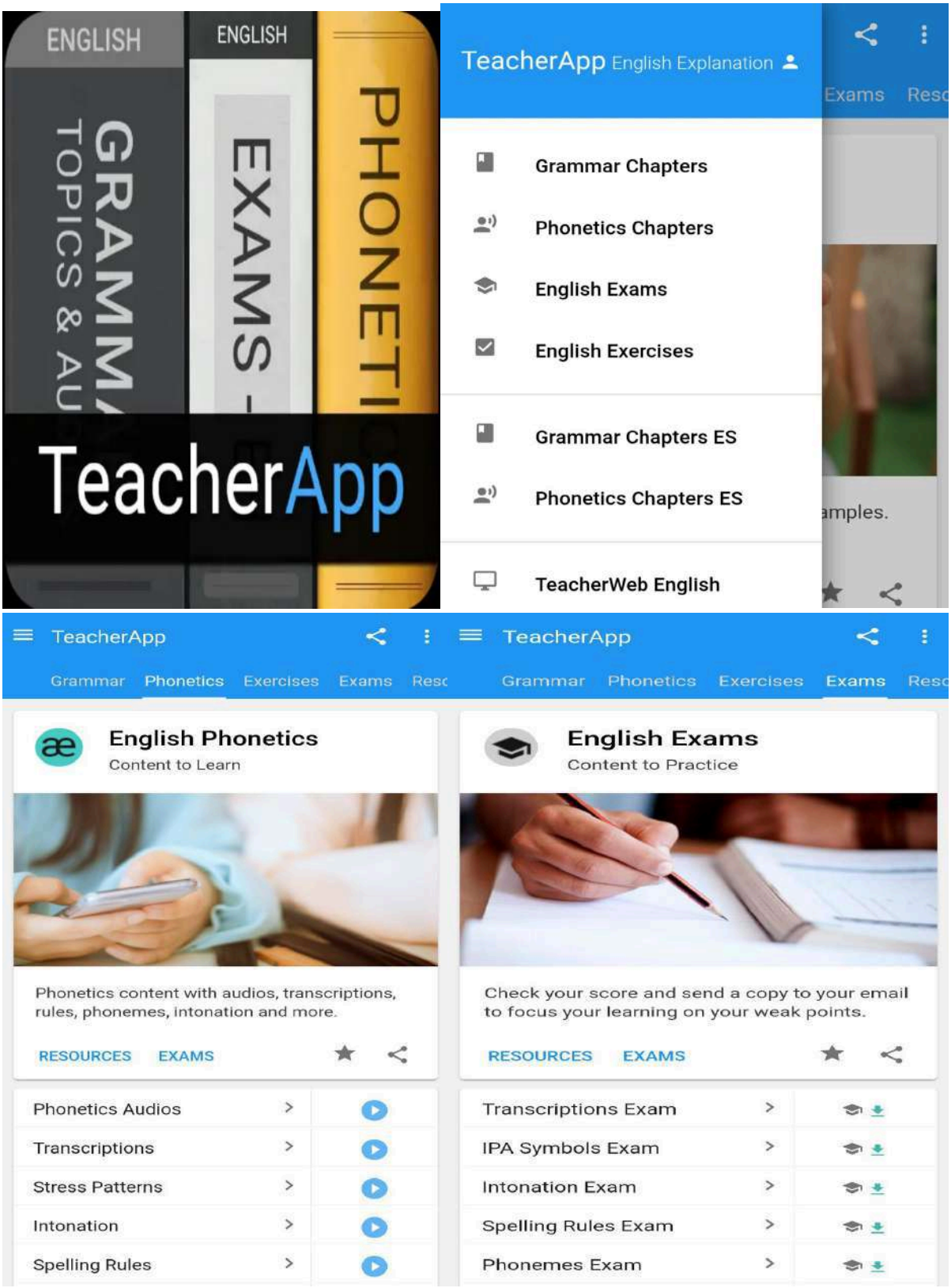




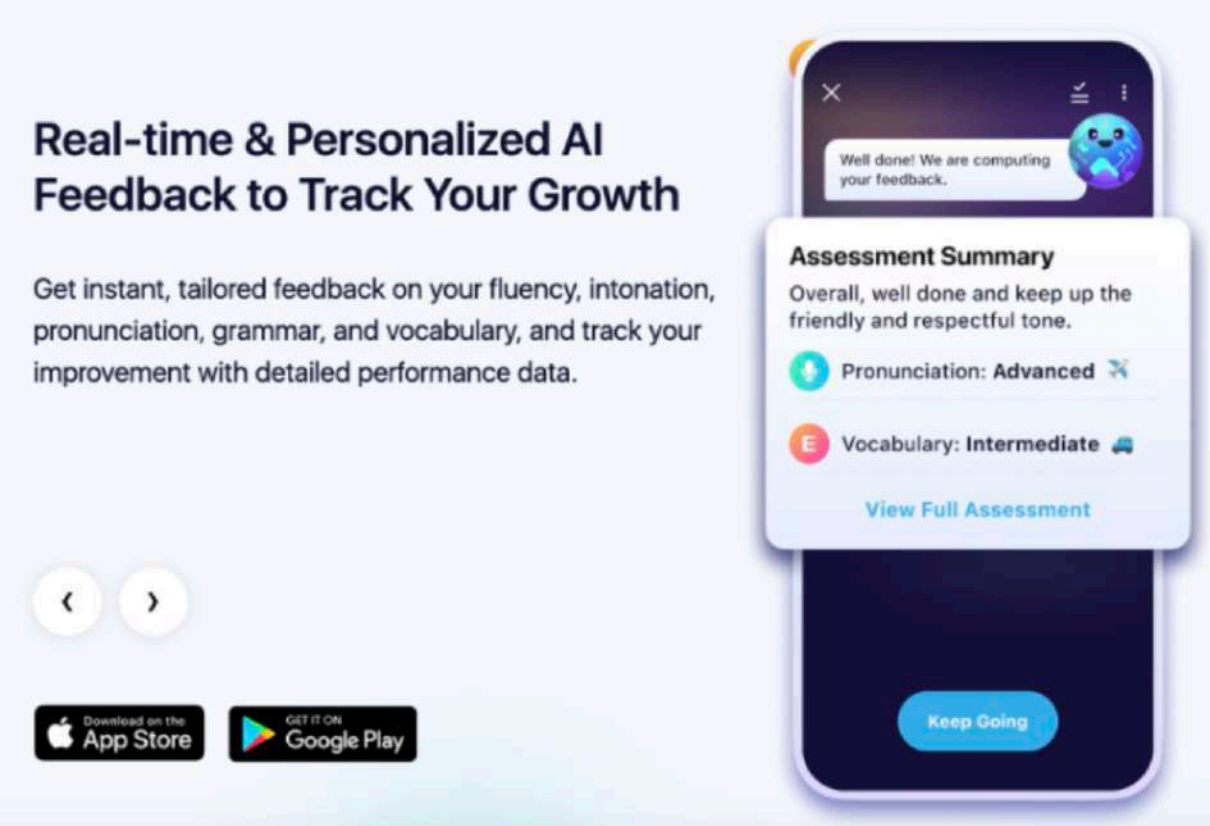


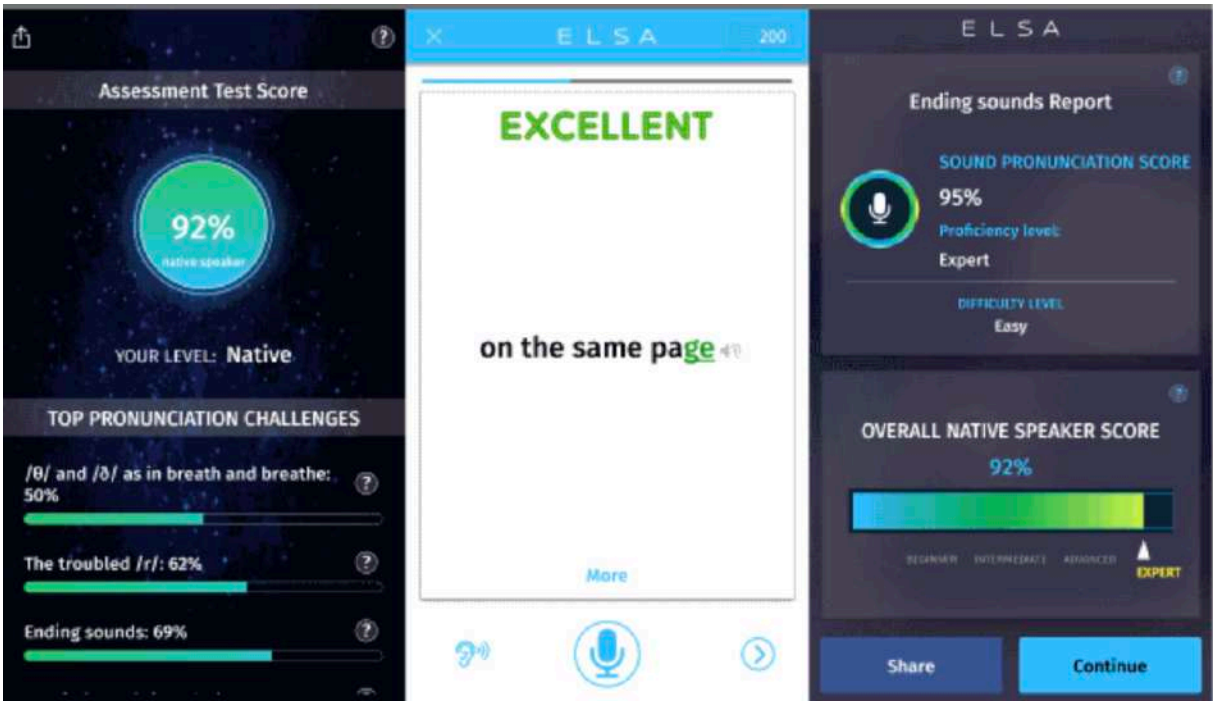

sanako
Menu
Help
SANAKO Lab 100 - Version 9.0
UNIVERSITY MOHAMED KHEIDER OF BISKRA
SANAKO Lab 100 Configuration
Find workstations
Set video sources
Set Audio Sources
Set video connections
Set Default Settings
Save
Exit
Teacher

Appendix 5: TeacherApp Mobile Application for Segmental Pronunciation Training



Appendix 6: ELSA Speak Mobile Application for Pronunciation Training and Feedback





Appendix 7: Pronunciation Training for Participants based on Multimedia

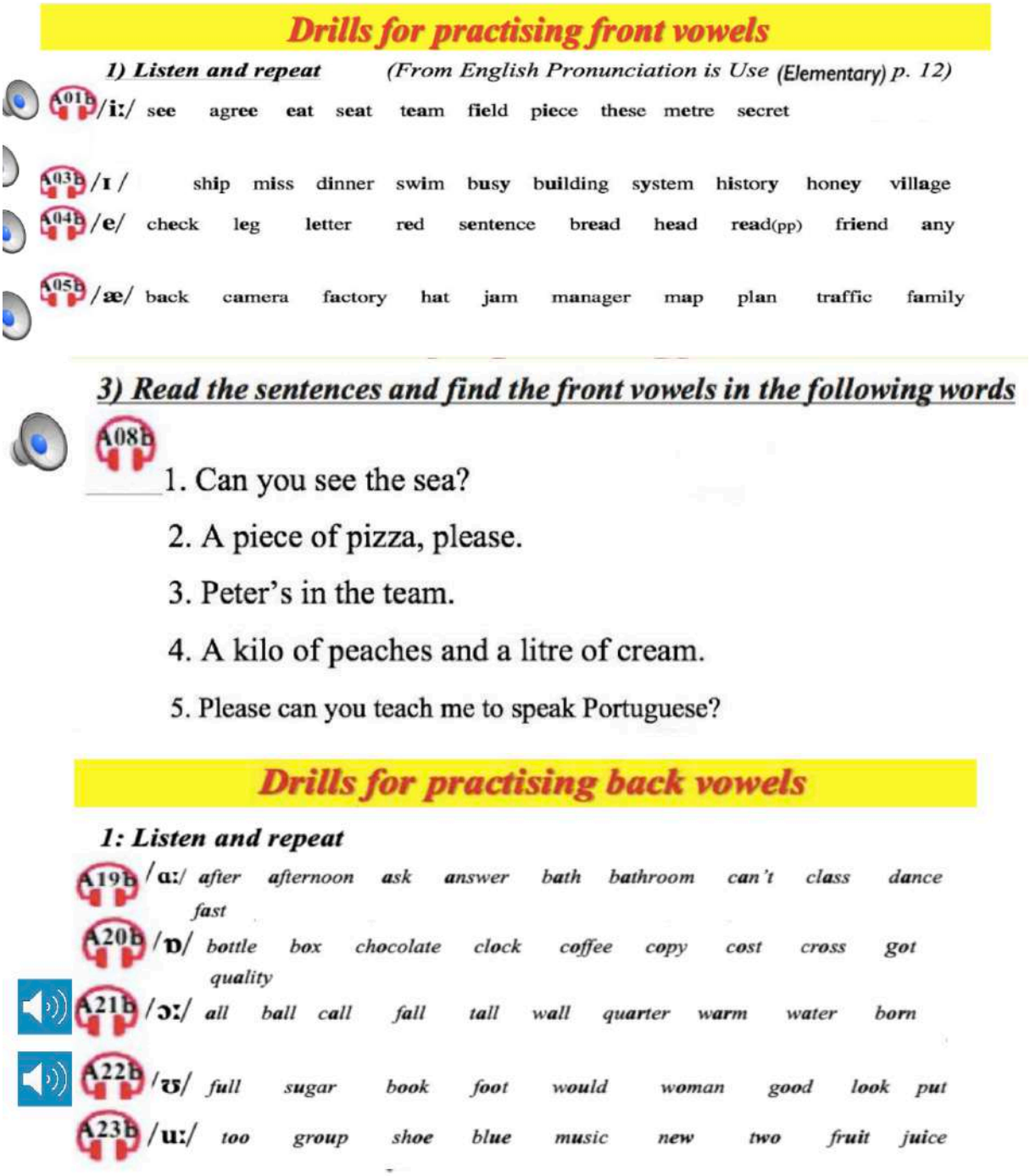
Listen to the Audio recordings and watch the videos of native-speakers

Task 1: Drills for practicing consonantal phonemes (from BBC Learning English)

Read aloud the words after listening to the audio

Read aloud the words after listening to the video

Task 2: Drills for practicing vowel phonemes in words and sentences (excerpted from English Pronunciation in Use (Elementary Level)



Drills for practising front vowels

1) Listen and repeat *(From English Pronunciation is Use* (Elementary) *p. 12)*

A01b /iː/ see agree eat seat team field piece these metre secret

A03b /ɪ / ship miss dinner swim busy building system history honey village

A04b /e/ check leg letter red sentence bread head read(pp) friend any

A05b /æ/ back camera factory hat jam manager map plan traffic family

3) Read the sentences and find the front vowels in the following words

A08b

1. Can you see the sea?
2. A piece of pizza, please.
3. Peter's in the team.
4. A kilo of peaches and a litre of cream.
5. Please can you teach me to speak Portuguese?

Drills for practising back vowels

1: Listen and repeat

A19b /ɑː/ *after afternoon ask answer bath bathroom can't class dance fast*

A20b /ɒ/ *bottle box chocolate clock coffee copy cost cross got quality*

A21b /ɔː/ *all ball call fall tall wall quarter warm water born*

A22b /ʊ/ *full sugar book foot would woman good look put*

A23b /uː/ *too group shoe blue music new two fruit juice*

Task 3: Drills for practicing word stress and intonation

(excerpted from English Pronunciation Programme)

Stress in Two- and Three-Syllable Words

Stress on the 1st Syllable	Stress on the 2nd Syllable	Stress on the 3rd Syllable
apple	invite	afternoon
teacher	tomorrow	absolute

Declarative sentences and requests end with a falling intonation. ↘

Claire is my friend. ↘
I bought a new car. ↘
Send the letter by airmail. ↘

Intonation in 'YES/NO' Questions

Questions that request a simple 'yes' or 'no' answer end with an upward intonation. ↗

May I see the menu? ↗
Does Michael smoke? ↗
Do you like the wine? ↗

Intonation in Questions/Answers

Be sure to use a rising pitch at the end of each yes/no question and a falling pitch at the end of the answer.

Will you go? ↗	Yes, I will. ↘
Did Paul study? ↗	No, he didn't. ↘

Noun Phrases ('s' = /s/)	**Verb Phrases** ('s' = /z/)
What's the use?	I use it.
Make an excuse.	Please excuse me.
I have a house.	The zoo will house the animals.

Appendix 8: Pronunciation Training using “the Chaos” Poem by Gerard Nolst Trenité

Dearest creature in creation,	/ˈdɪərɪst 'kri:ʧə ɪn kri:'eɪʃən /
Study English pronunciation.	/'stʌdi 'ɪŋglɪʃ prəˌnʌnsi:'eɪʃən/
I will teach you in my verse	/ɑɪ wɪl 'ti:ʧ ju: ɪn mɑɪ 'vɜ:s/
Sounds like corpse, corps, horse, and worse.	/saʊndz lɑɪk kɔ:ps kɔ: hɔ:s ənd wɜ:s/
Just compare heart, beard, and heard,	/ʤʌst kəm'peə hɑ:t bɪəd ənd hɜ:d/
Dies and diet, lord and word,	/dɑɪz ənd dɑɪət lɔ:d ənd wɜ:d/
Sword and sward, retain and Britain.	/sɔ:d ən swɔ:d rɪ'teɪn ənd 'brɪtn̩ /
Say break and steak, but bleak and streak;	/seɪ breɪk ənd steɪk bʌt bli:k ənd streɪk/
Woven, oven, how and low,	/ˈwəʊvn̩ ʌvən haʊ ənd ləʊ /
Script, receipt, shoe, poem, and toe.	/ˈskrɪpt rɪ'si:t ʃu: 'pəʊɪm təʊ /
Strewn with stones, stowed, solace, gunwale,	/stru:n wɪð 'stəʊnz stəʊd 'sɒləs 'gʌnəl/
Islington and Isle of Wight,	/'ɪzlɪŋtn ənd ɑɪl ɒv wɑɪt /
Saying lather, bather, father?	/'seɪɪŋ læðə beɪðə fɑ:ðə /
Finally, which rhymes with enough,	/'fɑɪnəli wɪʧ ˈrɑɪmz wɪð ɪ'nʌf /
Though, through, bough, or cough, hough or sough and tough?	/ðəʊ θru: bɑʊ kɒf hɒk sɑʊ ənd tʌf/
Hiccough has the sound of cup.	/ˈhɪkʌp həz ðə ˈsɑʊnd əv kʌp/
My advice is to give up!!!	/mɑɪ əd'vɑɪs ɪz tə gɪv ʌp/

Appendix 9: Summary of the Participants' Scores (Human Ratings for Pre-test)

Participants	***Segmental features scores***						***Suprasegmentals scores***		
(PreCG)	/i/	/ɪ/	/e/	/æ/	/ʊ/	/uː/	Minimal pairs	Stress	Intonation
Student 1	3	2	3	3	3	4	3	2	3
Student 2	2	1	2	3	4	4	3	3	2
Student 3	3	3	3	2	3	4	3	3	2
Student 4	4	3	2	3	4	3	3	2	3
Student 5	3	2	3	4	4	4	2	3	3
Student 6	2	3	3	3	3	4	3	3	2
Student 7	3	2	2	3	4	2	3	2	3
Student 8	4	3	1	2	3	4	2	3	2
Student 9	3	2	2	3	4	4	3	2	3
Student 10	2	3	3	2	3	3	2	3	2
(PreExG)									
Student 11	3	2	3	2	3	2	2	2	2
Student 12	2	3	2	3	4	3	2	2	3
Student 13	3	2	3	2	4	3	2	2	2
Student 14	4	3	3	2	3	2	1	2	3
Student 15	3	2	2	2	4	4	2	3	2
Student 16	3	2	1	2	3	3	2	3	3
Student 17	2	2	2	3	4	2	3	2	2
Student 18	3	2	3	2	3	2	2	3	2
Student 19	3	2	3	2	2	3	2	2	2
Student 20	3	2	1	3	2	2	1	2	2

10: Summary of the Participants' Gains (Human Ratings for Post-test)

Participants	*Segmental features gains*						*Suprasegmentals gains*		
(PostCG)	/iː/	/ɪ/	/e/	/æ/	/ʊ/	/uː/	Minimal pairs	Stress	Intonation
Student 1	3	3	5	4	4	4	4	3	4
Student 2	3	4	4	5	5	4	4	4	3
Student 3	4	5	5	4	4	4	4	4	3
Student 4	4	4	6	5	5	4	5	5	4
Student 5	5	4	5	5	6	4	5	4	3
Student 6	4	5	4	4	4	4	4	4	4
Student 7	5	6	4	5	5	3	4	3	4
Student 8	6	5	4	3	4	4	5	3	3
Student 9	4	5	5	4	4	5	4	4	3
Student 10	3	4	4	3	4	4	3	4	3
(PostExG)									
Student 11	4	4	6	5	6	5	6	4	4
Student 12	3	5	5	5	6	4	5	5	4
Student 13	6	5	6	4	5	5	4	4	3
Student 14	5	5	6	5	5	4	5	6	4
Student 15	6	6	5	5	6	5	6	6	4
Student 16	5	5	4	5	5	4	5	4	4
Student 17	6	6	5	5	5	3	5	4	4
Student 18	7	6	4	4	4	4	5	5	3
Student 19	5	5	6	4	5	5	5	4	4
Student 20	4	5	5	5	4	4	4	4	3