OPEN ACCESS

Australian Journal of Applied Linguistics
ISSN 2209-0959
https://www.castledown.com/journals/ajal/

Australian Journal of Applied Linguistics, 9, 103085 (2026)
https://doi.org/10.29140/ajal.2026.103085

The effects of video-dubbing tasks on EFL students' pronunciation of suprasegmental features and autonomy

MARWA SAID MUSTAFA EL-GARAWANY

Faculty of Archaeology and Languages, Matrouh University, Egypt

Abstract

The current study explored the effects of video-dubbing tasks on EFL students' pronunciation of suprasegmental features and autonomy. Participants (N = 10) were first-year students in the English Language Department, Faculty of Archaeology and Languages, Matrouh University. An EFL suprasegmental features pronunciation test and a learner autonomy scale were devised and administered before and after experimentation which lasted for seven weeks as part of the Phonetics and Phonology course during the second semester of the 2024–2025 academic year. Throughout the experiment, the students in pairs selected videos from the internet, removed their soundtracks, rehearsed the monologues and dialogues in the videos, created alternative soundtracks by recording their voices, integrated the videos with their own soundtracks, submitted the dubbed videos for feedback, and presented live dubbing to the whole class. Semi-structured interviews were also conducted to determine the students' perceptions toward the utility of video-dubbing tasks. Results indicated that video-dubbing tasks had significant effects on the students' pronunciation of suprasegmental features (linking, stress, and intonation) and autonomy. Besides, the students expressed highly positive perceptions toward the usefulness of video-dubbing tasks. Such results suggest that video-dubbing tasks can be used to promote EFL students' pronunciation of suprasegmental features and autonomy.

Keywords: Video-dubbing tasks, pronunciation of suprasegmental features, learner autonomy

Introduction

Pronunciation undoubtedly constitutes one of the crucial aspects of oral communication as it impacts the intelligibility and comprehensibility of speech. Pronunciation inaccuracies can lead to significant disruptions in communication, often resulting in more severe consequences than

those caused by grammatical or lexical mistakes. Hence, teaching and learning pronunciation is increasingly acknowledged within the field of FL acquisition (Jarosz, 2019; Pennington & Rogerson-Revell, 2019). Pronunciation refers to how sounds are produced and perceived by listeners (Richards & Schmidt, 2010). It encompasses various features of the speech stream at two levels: segmental (micro) and suprasegmental (macro). The former includes the distinctive features of individual sounds (i.e., vowels and consonants), whereas the latter involves those features extending over sounds, syllables, words, and sentences such as linking, stress, and intonation (Burns & Claire, 2003; Luoma, 2004). Research on pronunciation development primarily focuses on segmental features which are considered more facile to instruct and acquire (Levis, 2005; Saito, 2014). However, several scholars argue that suprasegmental features have a greater impact on intelligibility and should be pedagogically prioritized (Burns & Claire, 2003; Derwing & Munro, 2009; Hahn, 2004). Besides, EFL students often struggle to monitor their suprasegmental errors (McCrocklin, 2016), and teachers face challenges providing sufficient practice opportunities, particularly in authentic contexts, due to workload and time pressure (Offerman & Olson, 2016). Consequently, students need activities and tools that can empower them to develop these features independently, fostering their autonomy and reducing reliance on teachers for their learning progress (McCrocklin, 2016).

Utilizing task-based language teaching (TBLT) can promote EFL suprasegmental features production (Brown, 2021; Gordon, 2021; Jung et al., 2017; Tiittanen, 2017) and learner autonomy (Aldosari & Alsager, 2023; Lee, 2016; Vieira, 2017). TBLT is an instructional approach that underscores "the use of communicative and interactive tasks as the central units for the planning and delivery of instruction". Such tasks are described as workplans that foster meaningful interaction, facilitate negotiation of meaning, enable students to acquire various language skills through authentic language use, and necessitate a definitive communicative outcome (Richards & Schmidt, 2010, p. 585). Embracing the capabilities of technology through providing enriched input from multimedia and encouraging language production via repeated practice (Chapelle, 2003), various scholars have integrated it into TBLT, examining the usefulness of tasks mediated by computer- and/or web-based technologies (Al-Rashidi, 2025; Chaisiri, 2023; Chen, 2021; González-Lloret, 2022; Vellanki & Bandu, 2021). These technology-mediated tasks leverage the chosen technology to accomplish objectives that are not feasible within traditional classroom environments, enabling students to use both language and technology in productive and innovative manners (González-Lloret & Ortega, 2014). In addition to the characteristics of classroom-based tasks, they are designed to be learner-centred as they address students' specific needs and wants, allowing them to effectively use their linguistic resources, technological applications, and digital skills (González-Lloret, 2015).

Video-dubbing tasks, wherein students record their voices over selected video segments (Burston, 2005), represent one type of such technology-mediated tasks. To illustrate, video-dubbing tasks (a) enable students to select video clips for the task and collectively discuss the merits of each clip, thus fostering negotiation of meaning, (b) engage students in rehearsed interactions that mirror real-world situations, (c) entail the development of listening, reading, and speaking skills, (d) involve students cognitively by inviting them to interpret the emotions expressed in verbal interactions and make plans for the creation of the dubbed video, and (e) require the submission of the final dubbed video and/or the execution of live dubbing as the culminating product. Furthermore, video-dubbing tasks utilize video technology in ways that are not possible within conventional classroom settings, allowing students to showcase their linguistic gains during the dubbing process and/or the live presentation, while concurrently enhancing their video-editing skills (Huang, 2022; Hwang et al., 2025). Besides, video-dubbing tasks positively affect students' emotional states, leading to increased willingness to

communicate (Chang & Yeh, 2024) and reduced FL anxiety (Chang & Yeh, 2024; Huang, 2022; Jiménez & Espinoza, 2025). Commonly completed in pairs or groups, they cultivate a stronger peer or group cohesion (Huang, 2022). Nonetheless, given that it is a relatively new technique particularly in the EFL context, the number of empirical studies examining the didactic effects of video-dubbing tasks remains limited. Most of them have primarily concentrated on speaking skills, without directly addressing suprasegmentals or targeting the aspect of learner autonomy. Thus, this study explored the effects of video-dubbing tasks on EFL students' pronunciation of suprasegmental features and autonomy alongside their perceptions toward them.

Literature Review

Suprasegmental Features of Pronunciation

According to Crystal (2008, p. 466), the term suprasegmental refers to "a vocal effect which extends over more than one sound segment in an utterance, such as a pitch, stress, or juncture pattern". Similarly, Genetti (2019, p. 652) defines suprasegmental features as "the phonetic properties that extend across multiple sounds" such as linking, stress, and intonation. Concerning linking, Richards and Schmidt (2010, p. 343) refer to it as "a process in continuous speech which connects the final sound of one word or syllable to the initial sound of the next". It aims to merge two words into a single unit without altering their original segmental identity (Alameen & Levis, 2015). It involves various forms, namely vowel-vowel, consonant-vowel, and consonant-consonant linking. In vowel-vowel linking, words concluding with a tense vowel followed by a word or syllable commencing with a vowel are typically connected with a glide. To elaborate, words ending with /i:/, /eɪ/, /aɪ/, and /ɔɪ/ followed by a word starting with a vowel are linked by a small /j/ as in *be able* while words ending with /u:/, /əʊ/, and /aʊ/ followed by a word starting with a vowel are linked by a small /w/ as in *blue ink*. In certain English varieties, a linking /r/ is inserted between words that finish and begin with a vowel, as exemplified by *media event* or *saw Ann*. In consonant-vowel linking, the last consonant of a word links to the first vowel of the subsequent word as in *back on* and *some of*. When a word or syllable concluding with a consonant cluster precedes a syllable commencing with a vowel, the last consonant of the connecting cluster is frequently articulated as part of the subsequent syllable, a process known as resyllabification. For instance, *left arm* is typically articulated as *lef tarm*. Consonant-consonant linking occurs when two identical consonants appear at word boundaries, resulting in their pronunciation as a single, slightly lengthened sound as in *miss Sarah* and *five views* (Alameen & Levis, 2015; Ashton & Shepherd, 2012; Burns & Claire, 2003; Richards & Schmidt, 2010).

Regarding stress, Crystal (2008, p. 454) delineates it as "the degree of force used in producing a syllable". For Richards and Schmidt (2010, p. 560), it is "the pronunciation of a syllable or word with more respiratory energy or muscular force than other syllables or words in the same utterance" as listeners perceive stressed syllables or words as being more prominent (i.e., louder, longer, and higher in pitch) than surrounding syllables or words (Genetti, 2019). It divides into two types: word stress and sentence stress. Firstly, word stress pertains to the place of stressed and unstressed syllables within a word, particularly a long one, where the syllable with the greatest prominence has the primary stress and the following stressed syllable the secondary stress (Richards & Schmidt, 2010). English uses stress to distinguish between some pairs of words with distinct meanings. For instance, the word *INsight* is stressed on the first syllable, whereas the word *inCITE* is stressed on the second one. Moreover, English has various noun and verb pairs with different stress locations. For instance, the nouns *CONvert* and *IMport* have initial stress, while the verbs *conVERT* and *imPORT* have stress on the second syllable. In English, unlike other languages, the place of the primary stress is not predictable since stress

can occur on almost any syllable within a word (Genetti, 2019). Putting the primary stress on the wrong syllable often causes miscomprehension (Richards, 2015). Secondly, sentence stress pertains to the place of stressed and unstressed words within a sentence. It primarily falls on content words that convey added information. For instance, in the sentence *He was leaving LONdon*, the primary stress falls on the initial syllable of the word *London* (Richards & Schmidt, 2010). Sentence stress not only highlights semantically important words, but it also shapes the rhythm of an utterance (Burns & Claire, 2003).

Relating to intonation, Burns and Claire (2003, p. 7) describe it as "the melody of the language". Richards (2015, p. 341) defines it as "the variations in pitch when speaking- that is, the raising and lowering of pitch during an utterance … to carry meaning over and above the meaning of the utterance itself". Similarly, Genetti (2019, p. 48) delineates it as "the changes in fundamental frequency (pitch) that occur during a phrase or an utterance … to mark differences in meaning that are not conveyed by segmental differences". For instance, English statements and wh-questions are commonly marked by a pitch fall, whereas English yes/no questions are frequently marked by a pitch rise. The sentence *So, you don't believe that'll work* can be articulated with either a falling or a rising intonation pattern, leading to distinct connotations. The pitch fall suggests that the speaker is summing up the viewpoint of the person he/she is talking to, while the pitch rise implies an inquiry into whether the listener believes something will work or not (Genetti, 2019; Richards, 2015). From a different perspective, intonation serves four main functions: attitudinal, accentual, grammatical, and discoursal, making it a highly complex area of phonetics. Firstly, the attitudinal function enables the speaker to convey his/her attitudes and feelings such as pleasure, anger, annoyance, interest, or boredom. Secondly, the accentual function, often called the focusing or informational function, indicates which syllables are prominent within a spoken statement and draws the listener's attention to them. That is, the place of tonic stress may be changed according to the speaker's intent to emphasize specific elements in his/her message. Thirdly, the grammatical function helps the listener to recognize the start and end of grammatical units, differentiate between statements and questions, and disambiguate ambiguous sentences. Finally, the discoursal function enables the listener to distinguish between newly introduced and previously established information through highlighting new utterances and de-emphasizing old ones. It also allows the speaker to manage the conversation flow, signaling his/her conclusion or readiness to take a turn (Roach, 2009; Wells, 2006).

The production of suprasegmental features can therefore be problematic for EFL students, especially for those whose L1, such as Arabic, does not encompass many of these aspects (Abdelreheem, 2023; El-Garawany, 2021). This necessitates the implementation of effective pedagogical methods to address the recurring challenges students face in acquiring EFL suprasegmentals (Hewings, 2004). As per the researcher's knowledge, minimal empirical attempts have been undertaken to enhance them, especially through exploiting technology. For example, computer-assisted pronunciation training (CAPT) applications have been utilized to improve students' pronunciation of suprasegmental features by providing individualized feedback, increasing practice opportunities, and recording and visualizing students' speech signals as waveforms, spectrograms, and pitch contours (Alshehri, 2025; Li, 2019; Wong, 2016). For instance, using applications like MyET (Almasifar & Heidari, 2024; Bozorgian & Shamsi, 2020) and Praat (Aramipoor & Gorjian 2018; El-Garawany, 2021) significantly enhanced students' stress, rhythm, and intonation. Employing video technologies, videogames (Miranda Chávez, 2022) and video-based shadowing (Van et al., 2024) were used to develop students' linking, stress, and intonation. Besides, utilizing artificial intelligence (AI) programs such as ELSA Speak significantly promoted students' suprasegmental features in autonomous learning contexts (Senowarsito & Ardini, 2023).

Learner Autonomy

According to the self-determination theory, autonomy refers to a person tendency to actively participate in determining his/her own actions. When it comes to language learning, learner autonomy means the capacity of a student to take charge of various aspects related to his/her language learning (Benson, 2013; Richards & Schmidt, 2010). For Mercer et al. (2012, p. 251), it describes "a readiness to take control of one's own learning, which involves the ability to act independently and in cooperation with others in the service of one's own learning purposes". In the context of the EFL classroom, learner autonomy is especially important because when students are responsible for taking decisions concerning the planning and implementation of their learning tasks, they become more engaged in the learning process, which in turn allows them to be more proficient in the short- and long-term. The result is autonomous learners, as compared to those who depend on others to make such decisions (Lenkaitis, 2019; Richards & Schmidt, 2010). Considering this, Little (2007) posits three interrelated principles to develop learner autonomy: learner involvement, learner reflection, and target language use. Learner involvement indicates students' active participation in setting goals, selecting materials and activities, and evaluating outcomes. Learner reflection entails students' explicit reflective intervention which promotes metacognitive processes, particularly in the form of teacher- or student-student dialogues. Target language use calls for students' intensive use of the FL during both in-class and out-of-class activities, particularly through group projects that support writing and speaking skills.

Significantly, learner autonomy is positively correlated with various academic and affective outcomes. Students with high autonomy showed greater language proficiency, goals and needs satisfaction, metacognition satisfaction, intrinsic motivation, willingness to communicate, self-efficacy, and enjoyment (Alrabai, 2021; Barin & Eyerci, 2021; Celyan, 2021; Kassem, 2021). To reap such benefits, learner autonomy should be cultivated by adopting autonomy-supportive pedagogy that enables students to assume control over the whole or part of the learning process (O'Donnell et al., 2025). Therefore, various empirical investigations have been conducted to foster it, specifically through technology-enhanced interventions such as employing digital storytelling (Kim, 2014), videoconferencing (Lenkaitis, 2019), flipped teaching (Ghufron & Nurdianingsih, 2019; Khosravani et al., 2020), and multimedia learning (Liu et al., 2020). Technology-mediated TBLT using VoiceThread and Audioboo (Lee, 2016) and data-driven learning using corpus-based activities (Vedantham, 2025) were also used to improve it. Furthermore, AI advances such as automatic speech recognition (ASR) applications (McCrocklin, 2016; Inceoglu, 2022), AI-powered writing assistants (Dizon & Gold, 2023; Estrella, 2025), and AI-supported virtual tutoring systems (Hong & Guo, 2025) significantly promoted it within structured language learning environments.

Video-Dubbing Tasks

Video dubbing, also called lip-synchronized dubbing, entails substituting the original audio of a complete video or a nominated video clip with the voices of students and subdivides into intralingual dubbing which represents the "voice repetition of the original spoken language" and interlingual dubbing which embodies the "oral translation of the original spoken language" (Lertola, 2019, pp. 2–3). Intralingual dubbing encompasses two forms: simple and creative dubbing. Simple dubbing replaces the initial soundtrack with students' voiceovers without altering its original content. This requires students to mimic the intonation, pronunciation, and pace of native English speakers, thereby developing their speaking proficiency. In contrast, creative dubbing overlays a muted video clip with a modified script that alters the storyline. This requires students to use new lexical items and grammatical structures appropriate to

the video context, hence facilitating their discovery learning (Burston, 2005). In the EFL context, intralingual simple dubbing is often the primary method employed to support students' oral development (Danan, 2010; Pamungkas, 2019). Video-dubbing tasks primarily involve the auditory modality, enabling students to listen to spoken language, and at the same time the visual elements of the videos help students to reinforce auditory comprehension and familiarize themselves with the pronunciation and intonation of the FL (Bobkina et al., 2020). Besides, the dubbing process requires students to actively reproduce spoken language, thereby enhancing their speaking abilities, including accuracy, fluency, and comprehensibility. To achieve this, it is essential for students to repeatedly engage with the dialogues, thereby allowing for further practice opportunities beyond the classroom setting (Yachi & Karimata, 2008). With the advent of various mobile applications for video editing and dubbing, students are now afforded the opportunity to practice their speaking skills in authentic contexts, free from time or space constraints (Lertola, 2019; Pamungkas, 2019). Students can select their preferred video clips and autonomously monitor and develop their oral performance by dubbing and re-dubbing as necessary to achieve optimal end results (Burston, 2005).

To the researcher's knowledge, empirical research exploring the pedagogical usefulness of video-dubbing tasks in the EFL classroom has been sparse. Most studies have focused on boosting students' speaking skills such as comprehensibility, fluency, language accuracy, and accentedness (He & Wasuntarasophit, 2015; Huang, 2022; Hwang et al., 2025; Jao et al., 2022; Talaván & Costal, 2017; Wang & Lee, 2024), where video-dubbing tasks were perceived as linguistically beneficial and enjoyable (Hwang et al., 2025; Talaván & Costal, 2017), with only a few specifically targeting improvements in pronunciation performance. Chiu (2012) compared the effects of video dubbing and traditional instruction on the pronunciation of 83 Taiwanese first-year university students who were divided into a dubbing group (n = 41) and a control group (n = 42) during their EFL conversation course. The video-dubbing task necessitated a 10-minute movie clip script practice, followed by a live dubbing in front of the entire class. The quantitative results of the given questionnaire revealed that the dubbing group significantly outclassed the control group in terms of learners' goals, language use, and dramatization. The qualitative findings of the interviews indicated that video-dubbing activities were effective in improving pronunciation, fluency, and intonation awareness. Similarly, Karimzadeh and Ghahroudi (2017) explored the effects of video dubbing versus conventional instruction on the pronunciation development of 24 Iranian intermediate students who were divided into two groups of 12 during a speaking course. Unlike the control group, the experimental group practiced dubbing a 20-minute movie clip. The results of the pre-post test revealed that the dubbing group significantly outperformed the control group in producing English sounds and intonation. Pamungkas (2019) examined the impact of video dubbing tasks on the pronunciation of 20 Indonesian high-school students from the tenth and eleventh grades over eight classes. The results of the pre-post test showed significant improvements in the students' pronunciation of consonants, vowels, and diphthongs. Anh (2023) investigated the effects of video dubbing on promoting 40 Vietnamese high-school students' pronunciation. After four sessions in which the participants were required to dub four videos, the data were gathered through pretest, posttest, and teacher's observation. The results revealed that video dubbing significantly improved the students' different pronunciation aspects including vowels, consonants, stress, intonation, and intelligibility, and that the students had positive attitudes toward the implementation of video dubbing in their pronunciation training. Phan et al. (2025) examined the effects of video dubbing within the cognitive apprenticeship framework on 30 Vietnamese sixth-grade learners' pronunciation and their attitudes toward this technique. After ten-week in-class training, the findings of the pre-post test showed significant improvement in students' consonants, vowels,

stress, and intonation. The results of a questionnaire and semi-structured interviews revealed positive attitudes toward the technique.

A close examination of these studies indicates that, firstly, there is not yet enough research to evaluate these results in relation to suprasegmental features, namely linking, stress, and intonation. Secondly, existing research has considered the effectiveness of video-dubbing tasks primarily in terms of linguistic gains and student perceptions, overlooking their influence on individual difference constructs such as learner autonomy. No studies have been conducted to demonstrate autonomy changes resulting from this instructional method. Thirdly, while limited investigations have explored EFL students' attitudes of video-dubbing tasks (Anh, 2023; Chiu, 2012; Hwang et al., 2025; Phan et al., 2025; Talaván & Costal, 2017), further studies are required to better understand their perceptions about the technique, particularly with its impact on the production of suprasegmental features and learner autonomy. Fourthly, these previous studies utilized video-dubbing tasks that resulted in either a dubbed video clip (Anh, 2023; Karimzadeh & Ghahroudi, 2017; Pamungkas, 2019; Phan et al., 2025) or a live dubbing presentation (Chiu, 2012) as the final product during pronunciation training sessions. None explored projects requiring both for pronunciation development, despite the distinct educational benefits each provides since the former enhances language practice outside the classroom and the latter offers opportunities for fostering public speaking skills (Burston, 2005). To fill these gaps, this study aimed to contribute to the field by offering empirical evidence on the efficacy of video-dubbing tasks in enhancing EFL students' pronunciation of suprasegmental features and autonomy together with their perceptions toward them, requiring both a dubbed video clip and a live dubbing performance as outcomes.

Theoretical Support for the Assigned Video-Dubbing Tasks

In addition to the aforementioned technology-mediated TBLT framework, the assigned video-dubbing tasks are conceptually aligned with the sociocultural and skill acquisition theories. The sociocultural theory posits that learning is a mediated process where students construct knowledge through social interactions and cultural tools, including technological resources (Vygotsky, 1978). Incorporating video-dubbing tasks in a collaborative manner provides students with ample practice opportunities to speak and evaluate their performance before submitting their dubbed videos, allowing them to develop their production of suprasegmental features in an interactive and stimulating learning environment (Huang, 2022; Hwang et al., 2025). The skill acquisition theory emphasizes that focused practice is essential to convert declarative knowledge of pronunciation rules into practical behaviours of articulating specific pronunciation features (i.e., procedural knowledge). By utilizing video-dubbing tasks, students engage in repeated rehearsal during which the proceduralization process becomes progressively seamless and reaches automatization demonstrated by quicker reaction times and fewer mistakes (DeKeyser, 2015). Moreover, video-dubbing tasks are framed within three other theoretical foundations, namely the input, interaction, and output hypotheses. Krashen's (1982) input hypothesis asserts that FL learners acquire language most effectively when they are exposed to abundant comprehensible input. During video-dubbing tasks, watching videos with scripts acts as a rich source of input where students receive spoken input from the audio and written input from the scripts at the same time, leading to better recall results (Cokely & Muñoz, 2019; Lasekan & Godoy Pena, 2025). Related to this, the interaction hypothesis (Long, 1991) highlights the importance of negotiating meaning among students during their video-dubbing interactions as it ensures that the FL input that students receive is comprehensible, enabling them to acquire new pronunciation features successfully. Moreover, Swain's (1985) output hypothesis claims that FL learners require not only comprehensible input but also comprehensible output

to facilitate their language acquisition. When students engage in producing language output, they sometimes recognize gaps in their linguistic knowledge, prompting them to take appropriate measures to address those gaps. Video-dubbing tasks allow students to produce language in a contextualized manner, reinforcing their pronunciation skills and encouraging self-monitoring and correction (Huang, 2022; Hwang et al., 2025).

Method

Context of the Problem

It has been observed that Egyptian students majoring in English at the researcher's faculty encounter various problems in accurately producing suprasegmental features during their oral practices, assignments, and exams. Most students demonstrated limited abilities to (a) link words effectively in connected speech, resulting in choppy speech, (b) stress the correct syllables in words, or change the placement of stress in various lexical items, (c) emphasize important words in sentences or utterances, and (d) use the different pitch patterns properly, leading to misinterpretation of the intended meaning. They also experienced low autonomy. These problems may be due to the inadequate teaching and assessment methods, the limited practice opportunities in authentic and autonomy-supportive contexts, the minimal exposure to the FL, the negative transfer effect of the L1, and the different phonological systems of both languages (Abdelreheem, 2023; El-Garawany, 2021; Elsayed et al., 2021; Galal-Eldin, 2023). In Arabic, the glottal stop is a prevalent feature and there are relatively few words that begin with a vowel. Thus, Arab students often exhibit resistance to the phonetic linking of a final consonant to an initial vowel in English (Swan & Smith, 2001). Egyptian students also resist changes produced from connected speech processes, specifically the intrusion of /j/, /w/, and /r/ as in *stay up* /steɪ j ʌp/, *so awful* /səʊ w ɔ:fəl/, and *law and order* /lɔ: r ən ɔ:də/ to link a final vowel to an initial vowel (Abdelreheem, 2023). Concerning Arabic word stress, it is predictable and regular, thereby causing difficulty in understanding the unpredictable and irregular nature of English word stress and how it changes meaning or word class. Primary stresses are more common in Arabic and "unstressed syllables are pronounced more clearly with neutral vowels but not swallowed as in English" (Swan & Smith, 2001, p. 198). Egyptian students often exhibit inappropriate word stress placement and an increased number of stressed syllables in their English pronunciation (Abdelreheem, 2023; El-Garawany, 2021). Still, Arab students may encounter fewer challenges with phrase and sentence rhythms because of their similarity in both languages. Intonation patterns in Arabic also bear similarities to those found in English. However, Arab students frequently utilize rising intonation instead of structural markers to indicate questions, suggestions, and offers, a phenomenon that extends to English (Swan & Smith, 2001). They also utilize rising intonation in tag questions, regardless of whether the speaker implies agreement or anticipates agreement from the listeners (Kharma & Hajjaj, 1997). Egyptian students often misapply the various intonation patterns at the end of their English utterances (El-Garawany, 2021). To record the problem, the researcher explored the perceptions of first-year students toward suprasegmental features through semi-structured interviews with 13 participants during the second semester of the 2023-2024 academic year. Suprasegmentals, identified as crucial yet challenging for pronunciation, were believed to be aspects where most students rated their performance as average or below average. Students also exhibited medium or low autonomy and rarely practiced suprasegmentals in authentic and autonomy-supportive conditions. Therefore, this study aimed to help students to develop their pronunciation of suprasegmental features and autonomy through video-dubbing tasks. Specifically, it addressed the following research questions:

RQ1: To what extent can video-dubbing tasks affect EFL students' pronunciation of suprasegmental features and autonomy?
RQ2: What are the EFL students' perceptions of video-dubbing tasks?

Design and Participants

This study utilized the quan→QUAL mixed-methods approach employing the one-group pretest-posttest quasi-experimental research design. The quan→QUAL mixed-methods design was used due to the limited sample size and the absence of a comparison group, with the qualitative component serving as the predominating paradigm following an initial quantitative phase (Johnson & Christensen, 2014). Quantitatively, one test and one scale were administered before and after the experiment to assess the effects of video-dubbing tasks on the study group's pronunciation of suprasegmental features and learner autonomy. Qualitatively, semi-structured interviews were held with the students to get a more comprehensive and in-depth understanding of their perceptions toward video-dubbing tasks, hence bolstering the credibility of the quantitative findings. The participants (N = 10, aged 18–19) were first-year students (four males and six females) in the English Language Department, Faculty of Archaeology and Languages, Matrouh University, during the second semester of the 2024–2025 academic year. All students were intermediate learners of English, as determined through the entrance exam, who had studied EFL for about 12 years before university enrollment and never encountered or practiced video-dubbing tasks in any of their English courses. Besides, they were computer literate and comparable in their L1 (i.e., Arabic, the official language of Egypt) and pre-tertiary education. Before the study commenced, all students gave their written consent after being informed about its purpose and procedures, emphasizing anonymity and confidentiality. They understood that they had the right to withdraw at any time without penalty and that they were required to submit the ordinary assignments upon withdrawal since the completion of out-of-class assignments constitutes part of the course's grade according to the faculty's regulations. All students voluntarily participated and allowed data collection and dissemination.

Instruments

Three instruments were devised. First, the EFL suprasegmental features pronunciation test (SFPT) was used as a pre-/post-test. In this study, prior test knowledge effects were minimal since the focus was on measuring pronunciation- a skill-based ability rather than a knowledge-based one, making it less influenced by item knowledge (Phan et al., 2025). The test included one question to assess the students' ability to accurately produce suprasegmental features (see Appendix A). The students were required to read aloud a passage (adopted from Celce-Murcia et al., 1996, p. 398) which contained sufficient statements and questions to render the test valid (Hewings, 2004). Regarded as a form of controlled pronunciation production, the read-aloud task was chosen as it enhances comparability and promotes score reliability (Hewings, 2004; Richards, 2015). The students were individually tested and audio recorded within a time limit of 5 minutes; three of which were given to students to familiarize themselves with the passage. To ensure objectivity, two lecturers (the researcher and another lecturer with the same teaching experience and qualifications) measured the students' EFL pronunciation of suprasegmental features and calculated the mean. They used the developed rubric which included three skills: linking, stress, and intonation. Each skill was measured on a five-point rating scale (1 = *poor*, 2 = *below average*, 3 = *average*, 4 = *above average*, and 5 = *excellent*), leading to a total score ranging from three to 15. To establish the content validity of the SFPT,

a jury of five Egyptian university professors of TEFL/applied linguistics were requested to review it in terms of clarity of its question, suitability to the students' academic level, and relevance to the objective it intended to measure. They indicated that the test could be an accurate measure of EFL pronunciation of suprasegmental features. To calculate its reliability, the SFPT was piloted on a group of 13 first-year English language majors- out of the study sample- at the end of the second semester of the academic year 2023–2024. Since two lecturers scored it, the inter-scorer reliability was executed where the Pearson's correlation coefficient computed was above 0.9, thus signifying high test reliability. Second, The EFL learner autonomy scale (LAS) was an 11-item measure (see Appendix B), on which students rated their level of agreement according to a five-point rating scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *neither agree or disagree*, 4 = *agree*, and 5 = *strongly agree*). Thus, its score range was from 11 to 55. It was adapted from McCrocklin's (2016) pronunciation learner autonomy survey, which was designed to assess changes in students' beliefs regarding their autonomy after incorporating ASR in their pronunciation training. All items on the LAS were normally coded, except for the items LA3, LA4, and LA5 which were reverse coded prior to summing the responses. Pilot testing was conducted on the same students of the SFPT pilot test where the Cronbach's Alpha coefficient computed was also above 0.9, thereby indicating high scale reliability.

Third, the semi-structured interview guide (see Appendix C) included four open-ended questions to determine the students' perceptions about the usefulness of video-dubbing tasks. Soliciting voluntary participation, all students agreed to be interviewed. They were afforded a comfortable environment to freely express their views and experiences, with the researcher attentively listening without directing the interviews in any specific direction. The interviews lasted from 10–15 minutes and were audio recorded and transcribed. Following Johnson and Christensen's (2014) guidelines, the transcripts were carefully read, meaningfully segmented, and repeatedly reviewed to assign broad categories (codes) from the students' responses. These categories were further evaluated to reveal subthemes, reflecting more detailed aspects of the data. This iterative procedure enabled a thorough examination of the qualitative data, assuring precise identification and interpretation of primary themes. To promote the reliability of the analysis, the inter-coder reliability was employed wherein two coders (the same scorers of the SFPT) independently analyzed the data, collaboratively compared their findings, and reconciled any discrepancies in their coding until consensus was reached. This approach facilitated the cross-validation of identified themes and reduced potential biases or subjective interpretations.

Experimental Procedures

The experiment of this study was conducted as part of the Phonetics and Phonology course offered to first-year students who met with the researcher for four hours of face-to-face teaching each week. Prior to the experiment on 9th March 2025, the researcher administered the SFPT to the students to determine their entry level of EFL pronunciation of suprasegmental features. In a successive 15-minute session, the LAS was distributed to assess their levels of learner autonomy. During the first week, a one-hour orientation session was dedicated to presenting video dubbing to the students, demonstrating its definition, importance, and procedures. This session also aimed at training them on using the scoring rubric to evaluate their performance (see Appendix A). The students also had the autonomy to select their preferred application, given the plethora of free video editing tools available. Due to its accessibility, user-friendly design, and diverse features (Abdullah & Arief Muhsin, 2024; Chang & Yeh, 2024; Phan et al., 2025), all students opted for CapCut to facilitate their video dubbing. Finally, a hands-on video-dubbing exercise was performed to familiarize them with the process and a Telegram chat

group was created for the study group to share their dubbed videos and work screenshots and to conduct peer evaluations. Earlier in the course, in accordance with the course objectives, the students had acquired knowledge of suprasegmental features including linking, word stress, sentence stress, and intonation through lecturing, which was subsequently complemented by some 'listening and repeating' exercises. Henceforward, during Weeks 2–7, the students were required to complete three video-dubbing tasks as after-class assignments, each spanned two weeks. For each task (see Figure 1), the students collaborated in pairs and began the process by watching various video clips from well-known films and animations available on the internet to select a 3-4-minute clip- since long excerpts induce frustration rather than an enjoyable experience (Yen, 2021)- to be submitted to the researcher for approval. As recommended by Yen (2021), the researcher made sure that all students' clips had roughly equal language difficulty and included adequate verbal exchanges and appropriate content. Upon obtaining approval, they repeatedly watched and listened to the chosen videos, observed how the characters uttered their lines, and identified the linking, stress, and intonation patterns used in the scripts. After that, they muted or removed the original soundtracks, rehearsed the monologues and dialogues of the characters in the videos to closely match the suprasegmentals used, created alternative soundtracks by recording their voices, compared their recordings with the original ones to revise their dubbing products, integrated the videos with their own generated soundtracks, submitted the dubbed videos for peer feedback, reviewed them using the scoring rubric by identifying the merits and areas requiring improvement, uploaded the final enhanced dubbed videos (see Figure 2 for some screenshots from the students' work), and presented live dubbing to the whole class, with the muted video clips being projected on the whiteboard. The researcher commented on the submitted videos and the live presentations and offered constructive feedback when necessary.

After seven weeks, the researcher post-tested the study group applying the same instruments and following the same procedures in addition to the semi-structured interview guide on 27th April 2025. Owing to the small sample size, the students' scores on the pre- and post-tests

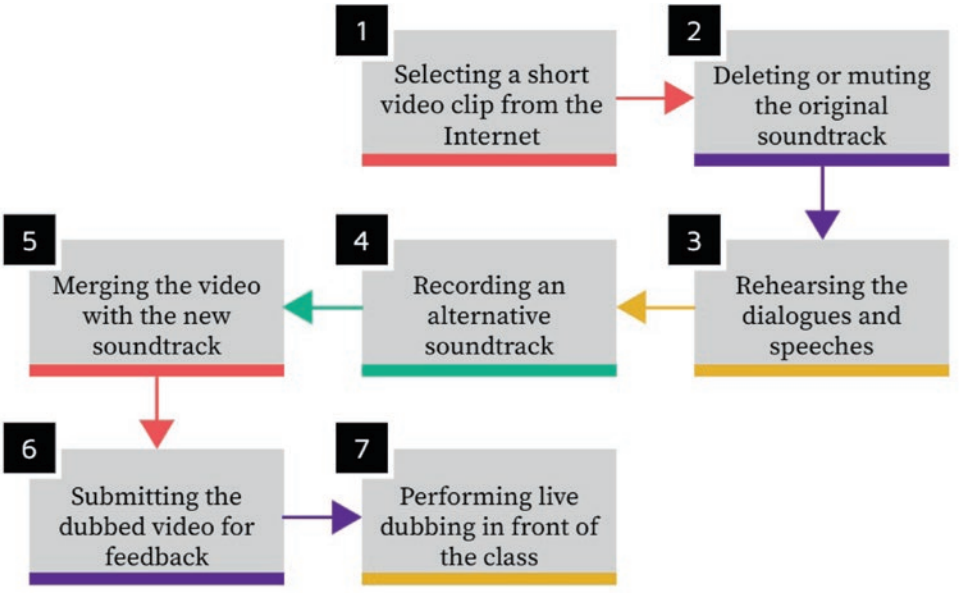

Figure 1. *Procedures of each video-dubbing task*

Figure 2. *Screenshots from the students' video-dubbing tasks on the CapCut application*

were analyzed using the nonparametric Wilcoxon Signed-Ranks Test. It was employed to compare the differences between the study group's mean ranks on both administrations to explore the effects of video-dubbing tasks on the two studied variables. Textual data from the students' semi-structured interviews were also analyzed using conventional content analysis to detect positive and negative themes regarding the utility of video-dubbing tasks.

Results

RQ1: To what extent can video-dubbing tasks affect EFL students' pronunciation of suprasegmental features and autonomy?

Table 1 displays the means, standard deviations, and *z*-values for the differences between the students' mean ranks of the EFL pronunciation of suprasegmental features and autonomy pre- and post-administrations. The mean values on the post-assessment of the EFL pronunciation of suprasegmental features and learner autonomy were higher than those on the pre-assessment, thus signifying the positive effects of the assigned video-dubbing tasks. The standard deviation values were generally small, denoting that the raw scores of the two main variables were comparatively consistent before and after the experiment. Table 1 also shows that there existed statistically significant differences at the 0.01 level between the students' mean ranks on the EFL pronunciation of suprasegmental features ($z = -2.8031$, $p < 0.01$)

Table 1. *Wilcoxon signed-ranks test results for the differences between the study group's mean ranks of the EFL pronunciation of suprasegmental features and learner autonomy pre- and post-administrations*

Dimension	Group	N	M	SD	Mean Rank	Sum of Ranks	z	Sig.	Effect Size
1. Linking	Pre	10	3.15	0.669	−0.85	55	−2.8031	0.01	0.9987
	Post	10	3.80	0.632					Large
2. Stress	Pre	10	3.30	0.587	−0.70	55	−2.8031	0.01	0.9429
	Post	10	3.85	0.580					Large
3. Intonation	Pre	10	3.25	0.677	−0.75	55	−2.8031	0.01	1.3375
	Post	10	4.15	0.669					Large
Suprasegmental Features (total)	Pre	10	9.70	1.874	−2.30	55	−2.8031	0.01	1.1297
	Post	10	11.80	1.844					Large
Learner Autonomy	Pre	10	31.40	2.797	−12.60	55	−2.8031	0.01	5.0571
	Post	10	43.20	1.751					Large

and autonomy (z = −2.8031, p < 0.01) pre- and post-administrations in favor of the post-administration. Thus, the study group made significant gains in EFL pronunciation of suprasegmental features and learner autonomy after employing video-dubbing tasks. Furthermore, the effect size was estimated using Cohen's d to determine the extent of the mean differences between the pre- and post-administrations for the two variables. The reported thresholds were set at 0.2 for small, 0.5 for moderate, and 0.8 for large effects. Table 1 indicates that the mean scores of both administrations in EFL pronunciation of suprasegmental features (d = 1.1297) and learner autonomy (d = 5.0571) differed significantly, as reflected by the substantial effect sizes. This might be attributed to the completion of the assigned video-dubbing tasks.

RQ2: What are the EFL students' perceptions of video-dubbing tasks?

The textual analysis of the students' semi-structured interviews generated seven positive themes and one negative theme as demonstrated in Table 2. To tangibly interpret the defined themes, some representative excerpts from the students were designated. For clarity and reference, these excerpts were labelled with students' identifiers (e.g., S1, S2, ..., S10). The ten students believed that completing their video-dubbing tasks helped them to enhance their pronunciation of suprasegmental features. They notably appreciated the different procedures employed in each video-dubbing task for enabling them to watch videos, analyze scripts, and engage in constant rehearsals to produce soundtracks that resembled the original ones. This helped them to accurately apply linking rules, stress assignment, and intonation patterns to produce natural and connected speech. They also indicated that implementing the required video-dubbing tasks as out-of-class assignments increased their autonomy. They had the freedom to select the short video clips and the editing application. They autonomously practiced and evaluated their suprasegmental features without solely relying on the lecturer for guidance.

Table 2. *Generated themes regarding the assigned video-dubbing tasks*

Theme	Number of Respondents	Representative Excerpts
Improved Pronunciation of Suprasegmental Features	10	– *Using video-dubbing tasks developed our suprasegmental features. We learnt to use the linking forms appropriately to make our speech sound more natural and connected. I also learnt that changing our intonation and tonic stress can change the meaning and impact of our utterances.* (S1) – *Watching videos, analyzing scripts, and engaging in repeated rehearsal to generate an alternative audio track resembling the original one improved our ability to apply linking rules, stress assignment, and intonation patterns properly.* (S2)
Increased Learner Autonomy	10	– *Completing the required video-dubbing tasks outside the classroom boosted our learner autonomy.* (S5) – *We were given the freedom to choose the films and the editing tool. We were able to practice suprasegmental features and correct our mistakes independently without relying entirely on our lecturer.* (S7)
Enhanced Enjoyment	8	– *I found the video-dubbing tasks more enjoyable than the traditional assignments. After submitting our dubbed videos on the Telegram group, our peers and you [the researcher] recommended some corrections. So, my partner and I edited our dubbed videos and submitted them again. We were very pleased with the positive feedback given after each resubmission and live performance.* (S1) – *I enjoyed the whole process so much especially when receiving positive comments on the dubbed videos or after the live dubbing which encouraged us to keep practicing and get improved.* (S3)
Heightened Motivation	8	– *We were delighted and motivated to continue practicing as a result of the progress we made with each video-dubbing task.* (S2) – *The fun we experienced during accomplishing each video-dubbing task and live performance encouraged us to persist in our practice.* (S9)

(Continue)

Table 2. *(Continued)*

Theme	Number of Respondents	Representative Excerpts
Benefits of Video-dubbing Tasks over Traditional Methods	7	– *I preferred video dubbing to listening and repeating because it enabled me to practice suprasegmentals, especially stress and intonation, repeatedly to reach the desired result using real-life situations from my favorite films.* (S7) – *I loved getting engaged in the video-dubbing tasks. This new technique was more useful than the traditional method because it made me understand and produce suprasegmental features in an interactive way.* (S8)
Facilitated Interaction and Collaboration	4	– *The assigned tasks provided us with opportunities to interact with our peers outside the classroom, allowing us to collaboratively analyze scripts, rehearse the dialogues, record audios, edit videos, and socialize*. (S6) – *We spent more time interacting and cooperating with each other to complete our tasks effectively*. (S7)
Increased Practice Opportunities	4	– *Using video-dubbing tasks as out-of-class assignments gave me more practice opportunities to improve my suprasegmental features in a less stressful and autonomy-supportive environment*. (S1) – *We engaged in repeated and focused practice to accomplish the assigned video-dubbing tasks on time*. (S9)
Technical Problems and Solutions	3	– *We sometimes had difficulty in merging the muted videos with the new recorded soundtracks. We needed to make several trials to generate the dubbed videos successfully*. (S4) – *We had some technical problems when we were uploading a modified part of our soundtracks after receiving feedback to match the lip synchronization and intonation of the characters. We had to make some adjustments to align the duration of the edited part with the original characters' speech and to properly insert it within the soundtracks*. (S10)

Eight participants thought that the assigned video-dubbing tasks presented a more enjoyable and motivating experience to improve their pronunciation of suprasegmentals. They recognized the entire process and valued the peer and teacher feedback received following the submission

of the dubbed videos. Based on this feedback, they continued to practice until their dubbed videos improved. They expressed satisfaction with the positive comments provided after each resubmission and live performance and reported that the enjoyment experienced and the progress attained over time fostered sustained engagement. Seven students preferred video-dubbing tasks to conventional methods for pronunciation training and found them more useful in enhancing their production of suprasegmental features and accomplishing their assignments in an interactive manner using authentic content and repeated practice. Four students stated that using video-dubbing tasks facilitated their social interactions and collaborative speaking as they worked together to analyze scripts, rehearse the dialogues, record audios, edit videos, and perform live dubbing for the whole class. Four students also noted that using video-dubbing tasks provided them with ample opportunities to practice suprasegmental features outside the classroom within a more relaxed and autonomy-supportive environment. Only three students reported some technical problems when merging the muted videos with the new soundtracks of their voices on the CapCut application and uploading modified segments after getting feedback to match the lip synchronization and intonation of the characters. For resolving these technical issues, numerous attempts were made to generate the dubbed videos successfully. Some adjustments were also made to align the duration of the edited sections with the original characters' speech and to appropriately integrate them into the soundtracks.

Discussion

Effects of Video-Dubbing Tasks

Regarding the first question, the quantitative findings showed that video-dubbing tasks had significant effects on the participants' pronunciation of suprasegmental features and autonomy. The significant improvement in their pronunciation of suprasegmental features might be ascribed to various reasons. Embracing the features of technology-mediated TBLT (González-Lloret, 2015; González-Lloret & Ortega, 2014), the assigned video-dubbing tasks allowed the students to practice pronunciation in authentic and meaningful contexts, which are often lacking in the traditional classroom, helping them to speak English more naturally as they imitated the characters' voices while maintaining the original speed. Guided by the sociocultural theory (Vygotsky, 1978) and the interaction hypothesis (Long, 1991), performing these tasks in pairs provided them with several opportunities to practice and evaluate their performance in an engaging learning environment to improve their suprasegmentals. The inclusion of peer feedback- using the developed rubric- on the submitted dubbed videos further enhanced the social dimension of the learning process, enabling students to communicate with each other and work together to identify their suprasegmental flaws and refine them. Drawing on the skill acquisition theory (DeKeyser, 2015), the repeated practice outside the classroom helped in improving the students' articulation speed and reducing their suprasegmental mistakes. The students engaged in the proceduralization process through systematic practice; they attentively analyzed the pronunciation patterns of native English speakers through listening and script reading, and through repeated speaking practice, they gradually converted their declarative knowledge into procedural knowledge to reach automatization. Along with practice, peer feedback might have contributed to the proceduralization and automatization processes as students got specific diagnostic feedback from their peers, facilitating their skill acquisition (El-Garawany, 2026; Sato & Lyster, 2012). Framed within the input (Krashen, 1982) and output (Swain, 1985) hypotheses, the frequent exposure to the target language might have assisted the students in producing accurate pronunciation as they realized gaps in their suprasegmental features when comparing their outputs with the characters' speech. To illustrate, the video visuals presented effective input for suprasegmental features by directing the students' attention

to these features. Such conscious attention assisted them in acquiring articulatory knowledge, correcting their own pronunciation, and producing improved pronunciation, perhaps preventing pronunciation fossilization (Wang & Lee, 2025). Since the voiceover task necessitates text-visual synchronization, ensuring linguistic and visual consistency (Talaván & Costal, 2017) and alignment with character lip movement (Chiu, 2012), the students observed and mimicked the lip movements of the characters. In this process, they focused on the facial expressions and emotions conveyed by the characters in the videos, integrating these aspects into their dubbing performances, which enhanced their tone, intonation, and authenticity of their speech. During the video-dubbing tasks, the students exerted considerable effort to improve their speaking rate, connect words, and sustain a continuous speech flow, which resulted in notable improvements in liaison and vocal intonation, including a heightened awareness of stress patterns and pitch variations. Moreover, the researcher's constructive feedback on the submitted videos and the live presentations might have contributed to the students' significant improvement in producing suprasegmental features.

Interestingly, the assigned video-dubbing tasks were beneficial in fostering the study group's overall pronunciation of suprasegmental features, as indicated by their scores on the post-assessment of the SFPT. Concerning linking, the students linked most words accurately and smoothly, with minimal errors, showing a good understanding of linking rules. Regarding stress, the students applied word and sentence stress correctly, with insignificant errors. As for intonation, which showed the greatest post-treatment improvement, the students used varied intonation patterns to convey the intended meanings and functions. One possible explanation for this might be the recurrence of listening and revoicing in contextualized situations while engaging in cognitive learning processes which required the students to focus on using the pitches, intonations, and other vocal attributes of the speech patterns (Chiu, 2011). Moreover, the audiovisual materials used might have offered substantial contextual clues that facilitated the students' comprehension of the language's natural intonational contours (Phan et al., 2025). This aligns with the results from studies by Chiu (2012), Karimzadeh and Ghahroudi (2017), Anh (2023), and Phan et al. (2025) who documented that video-dubbing tasks support EFL students' pronunciation of suprasegmental features (i.e., stress and intonation) and provide a valuable supplement to the traditional pronunciation methods.

Furthermore, video-dubbing tasks led to a significant increase in learner autonomy. The students' beliefs about their ability to manage their pronunciation learning were boosted. This might be attributed to the autonomy-supportive behaviours exhibited while executing their assigned tasks. They had the opportunity to collaborate on practicing and assessing their pronunciation performance of suprasegmental features outside the classroom, enabling them to choose when to engage with their video-dubbing tasks at their convenience, minimizing their reliance on the teacher feedback. They also had the freedom to select their preferred movie sequences and video-editing software, thus empowering them to have some control and influence their learning experience. This agrees with the findings from studies by Lee (2016), Vieira (2017), and Aldosari and Alsager (2023), who concluded that using TBLT and technology-mediated tasks promote EFL learner autonomy as they raise students' awareness of their own responsibility as learners.

Students' Perceptions of Video-Dubbing Tasks

Concerning the second question, the qualitative findings indicated that the participants displayed largely favorable perceptions toward the assigned tasks. Precisely, the number of positive excerpts was greater than that of negative ones, demonstrating that the study group highly appreciated these tasks in pronunciation and other facets of their language acquisition.

All students reported that the video-dubbing tasks were effective in improving their pronunciation of suprasegmental features (i.e., linking, stress, and intonation) and autonomy. As repeatedly disclosed by the interviews, this result might be accredited to the processes entailed by the completion of the given video-dubbing tasks outside the classroom, including viewing multiple films to choose suitable clips, analyzing scripts, attempting to reproduce the characters' pronunciation and intonation, and constantly rehearsing the selected dialogues to attain the desired effect. Given that videos constitute a genuine medium showcasing the behaviours of native speakers across diverse communicative contexts (Huang, 2022), this result might also be attributed to the students' exposure to authentic speech models which were emulated through performing the assigned video-dubbing tasks in autonomy-supportive settings, resulting in more accurate and natural speech outputs. The students also acknowledged video-dubbing tasks for being linguistically engaging and enjoyable. This result might be elucidated by highlighting Sánchez-Requena's (2016) claim that video-dubbing tasks are considered fun and motivating activities, encompassing role assumption, dialogue practice, and synchronization with on-screen characters. During these tasks, as asserted in the interviews, the students collaboratively rehearsed dialogues, synchronized their voice with on-screen characters, edited their dubbed videos based on the provided feedback, and performed live dubbing for the entire class. Besides, choosing popular films and/or animations might have increased student engagement in the learning process, since interaction with beloved characters and movies often leads to enhanced motivation (Phan et al., 2015). The students also preferred video-dubbing tasks to traditional methods in their pronunciation training. This result might have originated from the inherently appealing nature of the technique which is characterized by vibrant characters, images, and sounds, offering extensive practice opportunities to focus on certain aspects of pronunciation (Burston, 2005; Lertola, 2019). This enabled the students to repeatedly practice suprasegmental features till reaching the desired spoken outputs using real-life situations.

In addition, the students believed that video-dubbing tasks were helpful in facilitating their interaction and collaboration. This finding might have stemmed from the increased peer interaction opportunities outside the classroom, allowing the collaborative completion of necessary steps such as gathering around a computer for dubbing rehearsals (Burston, 2005), causing positive changes in peer bonding (Danan, 2010). This indication positions video-dubbing tasks as an effective communicative milieu for developing social strategies that are often associated with successful FL acquisition (Huang, 2022). Furthermore, it underscores the potential of video-dubbing tasks as collaborative learning activities that aligns with the principles of technology-mediated TBLT (González-Lloret, 2015; González-Lloret & Ortega, 2014). For example, these tasks stimulated negotiation of meaning through discussions about video selection and dubbing strategies, involving the participants in scripted interactions that resembled real-world scenarios. Moreover, the students stated that completing the allocated video-dubbing tasks provided them with ample practice opportunities beyond the classroom setting. This result could be ascribed to the necessity for participants to coordinate their utterances with filmic/animated characters during video dubbing. To achieve this, they had to focus on the characters' speech and rehearse more in a less stressful environment.

These qualitative results augment the quantitative findings and might elucidate why using video-dubbing tasks had positive effects on the students' pronunciation of suprasegmental features and autonomy. Still, a few technical problems were identified and resolved regarding merging the muted videos with the new recorded soundtracks and uploading modified parts of these soundtracks after receiving feedback to match the lip synchronization and intonation of the characters. These results agree with the findings by Chiu (2012), Anh (2023), and Phan et al.

(2025) who reported students' positive perceptions of video-dubbing tasks as being stimulating, motivating, and beneficial for their pronunciation development through mimicking the speech of native speakers despite some constraints.

Conclusion

This study aimed to examine the effects of video-dubbing tasks on EFL students' pronunciation of suprasegmental features and autonomy and determine the students' attitudes toward their utility. Three instruments were devised to collect data and address the formed questions. The results revealed that video-dubbing tasks significantly enhanced the students' pronunciation of suprasegmental features and autonomy. Besides, the students' perceptions toward them were truly positive. Thus, video-dubbing tasks proved to have positive effects on the study group's pronunciation of suprasegmental features and autonomy. In this regard, video-dubbing tasks should be considered for pronunciation training, offering students numerous opportunities to practice accurate pronunciation naturally and equipping university teachers, practitioners, and curriculum planners with innovative and creative instructional methods. While these results are promising, the absence of a control group, due to the limited enrollment in the English Language Department at the time of experimentation, raises questions about whether the assigned video-dubbing tasks were the primary factor that led to such positive findings. Moreover, the small homogeneous cohort of Egyptian English language majors engaged in this study might restrict the generalizability of the findings. In addition, this study was conducted over a short duration of seven weeks to manage maturation effects often associated with the one-group pretest-posttest design. Therefore, it is suggested that future research involve a control group and a larger sample from different EFL contexts over an entire academic semester or year. Other recommendations include replicating this study using other video-editing tools or investigating the effects of video-dubbing tasks on communication apprehension and speaking self-efficacy. In addition, subsequent studies can examine the impacts of other emerging educational technologies- such as AI, virtual reality (VR), digital storytelling, and gamification- on the acquisition of suprasegmental features. Given that the study of Phonetics and Phonology is frequently regarded as an abstract, difficult, and boring intellectual task, incorporating digital modalities that diversify instructional approaches may produce transformative benefits for student learning outcomes, engagement, and motivation.

Acknowledgements

The author gratefully acknowledges the first-year English majors who participated in this study and extends appreciation to the Editor-in-Chief, editorial team, and anonymous reviewers of the *Australian Journal of Applied Linguistics* for their careful reading and constructive feedback, which substantially improved this manuscript.

References

Abdelreheem, H. H. S. (2023). The key English pronunciation difficulties for Egyptian EFL learners. *Anglica: An International Journal of English Studies*, *32*(2), 115–136. https://doi.org/10.7311/0860-5734.32.2.06

Abdullah, N., & Arief Muhsin, M. (2024). How CapCut application complete video assignments: A study of students' perception in higher education in Indonesia. *Forum for University Scholars in Interdisciplinary Opportunities and Networking*, *1*(1), 600–610. https://conference.ut.ac.id/index.php/fusion/article/view/3345

Alameen, G., & Levis, J. M. (2015). Connected speech. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 157–174). John Wiley & Sons, Inc. http://dx.doi.org/10.1002/9781118346952.ch9

Aldosari, M. S., & Alsager, H. N. (2023). A step toward autonomy in education: Probing into the effects of practicing self-assessment, resilience, and creativity in task supported language learning. *BMC Psychology*, *11*, 1–15. https://doi.org/10.1186/s40359-023-01478-8

Almasifar, N., & Heidari, F. (2024). The effect of computer-assisted pronunciation training on EFL learners' use of suprasegmental features and foreign language speaking anxiety. *English Teaching & Learning*, *48*, 625–648. https://doi.org/10.1007/s42321-023-00159-4

Alrabai, F. (2021). The influence of autonomy-supportive teaching on EFL students' classroom autonomy: An experimental intervention. *Frontiers in Psychology*, *12*, 1–15. https://doi.org/10.3389/fpsyg.2021.728657

Al-Rashidi, A. H. (2025). Does digital task-based instruction make a difference in EFL university students' motivation in Saudi Arabia? An active theory perspective. *Learning and Motivation*, *90*, 102–115. https://doi.org/10.1016/j.lmot.2025.102115

Alshehri, A. A. (2025). The impact of computer-assisted pronunciation on Saudi students' production of English vowel sounds. *Australian Journal of Applied Linguistics*, *8*(1), 1–15. https://doi.org/10.29140/ajal.v8n1.102441

Anh, N. T. Y. (2023). The implementation of video dubbing in enhancing students' pronunciation at a high school in Vietnam. *International Journal of Social Science and Human Research*, *6*(2), 1254–1262. https://doi.org/10.47191/ijsshr/v6-i2-61

Aramipoor, S., & Gorjian, B. (2018). The effect of using PRAAT software on pre-intermediate EFL learners' supra segmental features. *Journal of Studies in Learning and Teaching English*, *7*(1), 47–68. https://sanad.iau.ir/en/Article/911557

Ashton, H., & Shepherd, S. (2012). *Work on your accent: Clearer pronunciation for better communication*. Collins.

Barin, M., & Eyerci, A. (2021). A survey on the relationship between willingness to communicate and learner autonomy of Turkish EFL students. *Shanlax International Journal of Education*, *9*(1), 104–116. https://doi.org/10.34293/education.v9iS1-May.4006

Benson, P. (2013). Learner autonomy. *TESOL Quarterly, 47*(4), 839–843. https://doi.org/10.1002/tesq.134

Bobkina, J., Domínguez Romero, E., & Gómez Ortiz, M. J. (2020). Educational mini-videos as teaching and learning tools for improving oral competence in EFL/ESL university students. *Teaching English with Technology*, *20*(3), 85–95. https://tewtjournal.org/download/6-educational-mini-videos-as-teaching-and-learning-tools-for-improving-oral-competence-in-efl-esl-university-students-by-jelena-bobkina-elena-dominguez-romero-and-maria-jose-gomez-ortiz/

Bozorgian, H., & Shamsi, E. (2020). Computer-assisted pronunciation training on Iranian EFL learners' use of suprasegmental features: A case study. *Computer-Assisted Language Learning Electronic Journal*, *21*(2), 93–113. https://old.callej.org/journal/21-2/Bozorgian-Shamsi2020.pdf

Brown, S. (2021). Teaching suprasegmentals within a task-based framework to improve speaking skills. *MinneTESOL Journal*, *37*(1), 1–8. https://h5j.a4b.myftpupload.com/wp-content/uploads/2021/05/Brown_Teaching-suprasegmentals-within-a-task-based-framework-to-improve-speaking-skills.pdf

Burns, A., & Claire, S. (2003). *Clearly speaking: Pronunciation in action for teachers*. AMEP Research Centre.

Burston, J. (2005). Video dubbing projects in the foreign language curriculum. *CALICO Journal*, *23*(1), 79–92. https://doi.org/10.1558/cj.v23i1.79-92

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.

Celyan, N. O. (2021). The relationship between learner autonomy and motivation. *TOJET: The Turkish Online Journal of Educational Technology*, *20*(4), 150–185. https://www.tojet.net/articles/v20i4/20414.pdf

Chaisiri, P. (2023). Promoting Thai EFL learners' willingness to communicate in the virtual classroom through technology-mediated oral tasks. *LEARN Journal: Language Education and Acquisition Research Network*, *16*(1), 97–120. https://so04.tci-thaijo.org/index.php/LEARN/article/view/263435

Chang, C., & Yeh, H.-C. (2024). Effects of video-dubbing creations (using peer instruction approaches) on learning achievement, willingness to communicate and learning anxiety. *The EuroCALL Review*, *31*(2), 17–33. http//doi.org/10.4995/eurocall.2024.21585

Chapelle, C. A. (2003). *English language learning and teaching*. John Benjamins.

Chen, K. T. (2021). The effects of technology-mediated TBLT on enhancing the speaking abilities of university students in a collaborative EFL learning environment. *Applied Linguistics Review*, *12*(2), 331–352. https://doi.org/10.1515/applirev-2018-0126

Chiu, Y. (2011). Can film dubbing projects facilitate EFL learners' acquisition of English pronunciation? *British Journal of Educational Technology, 43*(1), 24–27. http://dx.doi.org/10.1111/j.1467-8535.2011.01252.x

Cokely, M., & Muñoz, C. (2019). Captioned video, vocabulary and visual prompts: An exploratory study. In C. Herrero & I. Vanderschelden (Eds.), *Using film and media in the language classroom: Reflections on research-led teaching* (pp. 61–74). Multilingual Matters.

Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed.). Blackwell Publishing.

Danan, M. (2010). Dubbing projects for the language learner: A framework for integrating audiovisual translation into task-based instruction. *Computer Assisted Language Learning, 23*(5), 441–456. https://doi.org/10.1080/09588221.2010.522528

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed.) (pp. 94–112). Routledge.

Derwing, T., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, *42*(4), 476–490. https://doi.org/10.1017/S026144480800551X

Dizon, G., & Gold, J. (2023). Exploring the effects of Grammarly on EFL students' foreign language anxiety and learner autonomy. *The JALT CALL Journal*, *19*(3), 299–316. https://doi.org/10.29140/jaltcall.v19n3.1049

El-Garawany, M. S. M. (2021). Using Praat to develop English majors' EFL intonation production. *Sohag University Journal of Education*, *92*(3), 91–125. https://doi.org/10.21608/edusohag.2021.208708

El-Garawany, M. S. M. (2026). The effects of automatic speech recognition (ASR)-based peer feedback on EFL diphthong pronunciation improvement and retention. *The JALT CALL Journal*, *22*(1), 1–31. https://doi.org/10.29140/jaltcall.2026.102857

Elsayed, A. A. A., Mustafa, A., & Arbab, A. N. (2021). Pronunciation patterns of Omani EFL learners. *Journal of Basic and Environmental Sciences*, *8*(4), 166–170. https://doi.org/10.21608/jbes.2021.372054

Estrella, F. (2025). Supporting independent writers: How Grammarly shapes learner autonomy and writing performance in the Ecuadorian foreign language context. *Social Education Research*, *6*(1), 125–137. https://doi.org/10.37256/ser.6120255896

Galal-Eldin, N. (2023). Types of English pronunciation errors among Egyptian English major postgraduate students. *Journal of the Faculty of Arts- Fayoum University*, *15*(1), 1976–2030. https://doi.org/10.21608/jfafu.2023.239534.1964

Genetti, C. (2019). Syntax: Words in combination. In C. Genetti (Ed.), *How languages work: An introduction to language and linguistics* (2nd ed.) (pp. 129–136). Cambridge University Press.

Ghufron, M. A., & Nurdianingsih, F. (2019). Flipped teaching with CALL in EFL writing class: How does it work and affect learner autonomy? *European Journal of Educational Research*, *8*(4), 983–997. http://doi.org/10.12973/eu-jer.8.4.983

González-Lloret, M. (2015). *A practical guide to integrating technology into task-based language teaching*. Georgetown University Press.

González-Lloret, M. (2022). Technology-mediated tasks for the development of L2 pragmatics. *Language Teaching Research*, *26*(2), 173–189. https://doi.org/10.1177/13621688211064930

González-Lloret, M., & Ortega, L. (2014). *Technology-mediated TBLT: Researching technology and tasks*. John Benjamins Publishing Company.

Gordon, J. (2021). Pronunciation and task-based instruction: Effects of a classroom intervention. *RELC Journal*, *52*(1), 94–109. https://doi.org/10.1177/0033688220986919

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201–203. https://doi.org/10.2307/3588378

He, P., & Wasuntarasophit, S. (2015). The effects of video dubbing tasks on reinforcing oral proficiency for Chinese vocational college students. *Asian EFL Journal*, *17*(2), 106–133. https://repo.uinmybatusangkar.ac.id/xmlui/bitstream/handle/123456789/8480/1509097091385_15263591900%20Jun%2015.pdf?sequence=1&isAllowed=y

Hewings, M. (2004). *Pronunciation practice activities: A resource book for teaching English pronunciation*. Cambridge University Press.

Hong, X., & Guo, L. (2025). Effects of AI-enhanced multi-display language teaching systems on learning motivation, cognitive load management, and learner autonomy. *Education and Information* Technologies, *30*, 17155–17189. https://doi.org/10.1007/s10639-025-13472-1

Huang, H. D. (2022). Investigating the influence of video-dubbing tasks on EFL learning. *Language Learning & Technology*, *26*(1), 1–20. https://hdl.handle.net/10125/73489

Hwang, G.-J., Fathi, J., & Rahimi, M. (2025). Fostering EFL learners' speaking skills and flow experience with video-dubbing tasks: A flow theory perspective. *Journal of Computer Assisted Learning*, *41*(2), 1–18. https://doi.org/10.1111/jcal.13120

Inceoglu, S. (2022). Developing pronunciation learner autonomy with automatic speech recognition and shadowing. In S. McCrocklin (Ed.), *Technological resources for second language pronunciation learning and teaching: Research-based approaches* (pp. 171–192). Lexington Books.

Jao, C.-Y., Yeh, H.-C., Huang, W.-R., & Chen, N.-S. (2022). Using video dubbing to foster college students' English-speaking ability. *Computer Assisted Language Learning*, *37*(4), 585–607. https://doi.org/10.1080/09588221.2022.2049824

Jarosz, A. (2019). *English pronunciation in L2 instruction: The case of secondary school learners*. Springer Nature Switzerland.

Jiménez, T. F., & Espinoza, A. Q. (2025). Video dubbing as a strategy for reducing foreign language speaking anxiety levels. *Revista Innovaciones Educativas*, *27*(42), 29–41. https://doi.org/10.22458/ie.v27i42.5301

Johnson, R. B., & Christensen, L. (2014). *Educational research: Quantitative, qualitative, and mixed approaches* (5th ed.). Sage Publications.

Jung, Y., Kim, Y., & Murphy, J. (2017). The role of task repetition in learning word-stress patterns through auditory priming tasks. *Studies in Second Language Acquisition*, *39*(2), 319–346. https://www.jstor.org/stable/26330761

Karimzadeh, P., & Ghahroudi, M. R. (2017). English animation dubbing based techniques and Iranian intermediate EFL learners' nativelike pronunciation development. *International Journal of English Language Education*, *5*(2), 1–12. https://doi.org/10.5296/ijele.v5i2.11175

Kassem, H. M. (2021). EFL enjoyment of Saudi English majors and its relation to their achievement, autonomy, and self-Efficacy. *Journal of Education Sohag University*, *82*(2), 61–90. https://doi.org/10.21608/edusohag.2021.137623

Kharma, N., & Hajjaj, A. (1997). *Errors in English among Arabic speakers: Analysis and remedy*. York Press.

Khosravani, M., Khoshsima, H., & Mohamadian, A. (2020). On the effect of flipped classroom on learners' achievement, autonomy, motivation and WTC: Investigating learning and learner variables. *Journal of English Language Teaching and Learning*, *12*(25), 175–189. https://doi.org/10.22034/elt.2020.10680

Kim, S. (2014). Developing autonomous learning for oral proficiency using digital storytelling. *Language Learning & Technology*, *18*(2), 20–35. http://dx.doi.org/10125/44364

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon.

Lasekan, O., & Godoy Pena, M. T. (2025). Gender and academic discipline influences on EFL learners' perceptions of video-based instruction. *Australian Journal of Applied Linguistics*, *8*(3), 1–20. https://doi.org/10.29140/ajal.v8n3.2185

Lee, L. (2016). Autonomous learning through task-based instruction in fully online language courses. *Language Learning & Technology*, *20*(2), 81–97. http://dx.doi.org/10125/44462

Lenkaitis, C. (2019). Technology as a mediating tool: Videoconferencing, L2 learning, and learner autonomy. *Computer-assisted Language Learning*, *33*(5-6), 483–509. https://doi.org/10.1080/09588221.2019.1572018

Lertola, J. (2019). *Audiovisual translation in the foreign language classroom: Applications in the teaching of English and other foreign languages*. Research-publishing.net.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*(3), 369–377.

Li, K. (2019). The application of Praat in English pronunciation teaching. *Advances in Social Science, Education, and Humanities Research*, *378*, 374–376. https://dx.doi.org/10.2991/assehr.k.191217.150

Little, D. (2007). Language learner autonomy: Some fundamental considerations revisited. *Innovation in Language Learning and Teaching*, *1*(1), 14–29. https://doi.org/10.2167/illt040.0

Liu, X., Liu, Y., & Tu, J.-F. (2020). Multimedia technology and learner autonomy: An experimental study for asymmetric effects. *Symmetry*, *12*(3), 1–11. https://doi.org/10.3390/sym12030462

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of research on second language acquisition* (pp. 413–468). Academic Press.

Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of automatic speech recognition. *System*, *57*(2016), 25–42. http://dx.doi.org/10.1016/j.system.2015.12.013

Mercer, S., Ryan, S., & Williams, M. (2012). *Psychology for language learning: Insights from research, theory and practice*. Palgrave Macmillan.

Miranda Chávez, F. F. (2022). *Videogames and suprasegmental features of English*. https://repositorio.uta.edu.ec/items/a4f6e085-6ab0-4e82-b007-ddc0505ecf61

O'Donnell, A. M., Dobozy, E., Nagel, M. C., Bartlett, B., Smala, S., Wormald, C., Yates, G., Reeve, J., & Smith, J. K. (2025). *Educational psychology* (4th ed.). Wiley.

Offerman, H. M., & Olson, D. J. (2016). Visual feedback and second language segmental production: The generalizability of pronunciation gains. *System*, *59*(2016), 45–60. https://doi.org/10.1016/j.system.2016.03.003

Pamungkas, F. D. (2019). How video dubbing can improve students' speaking pronunciation. *ETERNAL (English, Teaching, Learning, and Research Journal)*, *5*(1), 41–53. https://doi.org/10.24252/Eternal.V51.2019.A4

Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: Contemporary perspectives*. Palgrave Macmillan.

Phan, V. T. T., Nguyen, T. Q., & Nguyen, K. D. (2025). The effects of video dubbing on EFL learners' pronunciation and their attitudes toward the technique. *TESL-EJ*, *29*(1), 1–35. https://doi.org/10.55593/ej.29113a4

Richards, J. C. (2015). *Key issues in language teaching*. Cambridge University Press.

Richards, J., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Pearson Education Limited.

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

Sánchez-Requena, A. (2016). Audiovisual translation in teaching foreign languages: Contributions of dubbing to develop fluency and pronunciation in spontaneous conversations. *Porta Linguarum*, *26*(June), 9–21. https://digibug.ugr.es/bitstream/handle/10481/53920/ART1_Alicia%20Sanchez.pdf?sequence=1&isAllowed=y

Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, *34*(4), 591–626. https://doi.org/10.1017/S0272263112000356

Senowarsito, S., & Ardini, S. N. (2023). The use of artificial intelligence to promote autonomous pronunciation learning: Segmental and suprasegmental features perspective. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *8*(2), 133-147. http://dx.doi.org/10.21093/ijeltal.v8i2.1452

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.

Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press.

Talaván, N., & Costal, T. (2017). iDub-the potential of intralingual dubbing in foreign language learning: How to assess the task. *Language Value*, *9*(1), 62-88. https://doi.org/10.6035/languagev.2017.9.4

Tiittanen, M. (2017). The teaching of suprasegmentals within a task-based or communicative ESL class. *CONTACT Magazine*, *August* (2017), 37–40. https://contact.teslontario.org/wp-content/uploads/2017/08/Tittanen-Suprasegmentals.pdf

Van, P. T. T., Lien, L. K., & Khang, N. D. (2024). Effects of video-based shadowing on suprasegmental features: EFL learners' pronunciation performance and attitudes. *rEFLections*, *31*(3), 896–924. https://so05.tci-thaijo.org/index.php/reflections/article/view/275498/183671

Vedantham, R. (2025). The influence of data-driven learning on learner autonomy and vocabulary acquisition. *Research Studies in English Language Teaching and Learning*, *3*(2), 398–410. https://doi.org/10.62583/rseltl.v3i2.83

Vellanki S. S., & Bandu, S. (2021). Engaging students online with technology-mediated task-based language teaching. *Arab World English Journal (AWEJ)*, *Special Issue on Covid 19 Challenges* (2021), 107–126. https://dx.doi.org/10.24093/awej/covid.8

Vieira, F. (2017). Task-based instruction for autonomy: Connections with contexts of practice, conceptions of teaching, and professional development strategies. *TESOL Quarterly*, *51*(3), 693–715. https://doi.org/10.1002/tesq.384

Vygotsky, L. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.

Wang, X., & Lee, S.-M. (2024). The impact of video dubbing app on Chinese college students' oral language skills across different proficiency levels. *International Journal of Educational Research*, *130*, 1–13. https://doi.org/10.1016/j.ijer.2024.102521

Wells, J. (2006). *English intonation: An introduction*. Cambridge University Press.

Wong, C. (2016). Teaching pronunciation to learners of English as a lingua franca. In W. Renandya & H. Widodo (Eds.), *English language teaching today: Linking theory and practice* (pp. 241–255). Springer International Publishing Switzerland.

Yachi, M., & Karimata, E. (2008). Online video dubbing system for language education. In C. Bonk, M. Lee, & T. Reynolds (Eds.), *Proceedings of e-learn 2008- World conference on e-learning in corporate, government, healthcare, and higher education* (pp. 4029–4034). Association for the Advancement of Computing in Education (AACE).

Yen, T. T. H. (2021). Using movie dubbing to improve natural English pronunciation skills. *English Teaching Forum*, *59*(1), 20–25. https://americanenglish.state.gov/files/ae/resource_files/etf_59_1_pg20-25.pdf

Appendix A
The EFL Suprasegmental Features Pronunciation Test (SFPT)

(Passage adopted from Celce-Murcia et al., 1996, p. 398)

– Read the following passage aloud: (15 marks)

If English is not your native language, people may have noticed that you come from another country because of your "foreign accent." Why do people usually have an accent when they speak a second language? Several theories address this issue. Many people believe that only young children can learn a second language without an accent, but applied linguists have reported cases of older individuals who have mastered a second language without an accent. Another common belief is that your first language influences your pronunciation in a second language. Most native speaker of English can, for example, recognize people from France by their French accents. They may also be able to identify Spanish or Arabic speakers over the telephone, just by listening carefully to their pronunciation. Does this mean that accents can't be changed? Not at all! But old habits won't change without a lot of hard work, will they? In the end, the path to learning to speak a second language without an accent appears to be a combination of hard work, a good ear, and a strong desire to sound like a native speaker. You also need accurate information about the English sound system and lots of exposure to the spoken language. Will you manage to make progress, or will you just give up? Only time will tell, I am afraid. Good luck and don't forget to work!

The Scoring Rubric for the EFL Suprasegmental Features Pronunciation Test (SFPT)

Level / Skill	1 Poor	2 Below Average	3 Average	4 Above Average	5 Excellent
1. Linking	Students are unable to link words accurately, causing fragmented speech, showing little or no understanding of linking forms.	Students link few words accurately, with many errors and unnatural pauses, showing a limited understanding of linking forms.	Students link some words accurately, with some noticeable errors and unnatural pauses, showing a basic understanding of linking forms.	Students link most words accurately and smoothly, with minimal errors, showing a good understanding of linking forms.	Students link words accurately and smoothly with perfect precision and consistency, with no unnatural pauses, showing an advanced understanding of linking forms.
2. Stress	Students completely misapply word and sentence stress, which causes significant problems in comprehension.	Students frequently misapply word and sentence stress, which interferes with comprehension.	Students fairly apply word and sentence stress, with some noticeable errors.	Students accurately apply word and sentence stress, with minimal errors.	Students expertly apply word and sentence stress, with perfect precision and natural rhythm.
3. Intonation	Students use intonation patterns inappropriately, which causes significant problems in comprehension.	Students use limited intonation variation, often monotone, which interferes with comprehension.	Students fairly use varied intonation patterns to convey the intended message, with some noticeable errors which mislead listeners.	Students accurately use varied intonation patterns to convey the intended meanings and functions, with minimal errors.	Students expertly use varied intonation patterns to convey the intended meanings and functions.

Appendix B
The Learner Autonomy Scale (LAS) Items

(adapted from McCrocklin, 2016)

LA1:	I am concerned about my English pronunciation.
LA2:	I want to continue improving my English pronunciation.
LA3:	I need a native speaker to learn how to pronounce English clearly and comprehensibly.
LA4:	I need a native speaker to correct my English pronunciation errors to improve.
LA5:	I need a teacher to assist me in improving my English pronunciation.
LA6:	I can use multiple pronunciation learning strategies effectively.
LA7:	I can successfully practice my English pronunciation on my own.
LA8:	I can successfully evaluate my English pronunciation on my own.
LA9:	I have the resources and tools that can help me to work on my English pronunciation.
LA10:	I can use technology to help me with my English pronunciation.
LA11:	I like trying new methods and activities to improve my English pronunciation.

Appendix C
The Semi-structured Interview Guide

1. Can you describe your experience of accomplishing each video-dubbing task? What features did you find most useful and what difficulties did you encounter during the video-dubbing process?
2. How do you think video-dubbing tasks impacted your EFL pronunciation of suprasegmental features and learner autonomy? Please, mention specific examples to illustrate this impact.
3. In what ways did video-dubbing tasks differ from conventional instruction? Which method do you consider to be more useful, and for what reason?
4. Did you enjoy accomplishing the assigned video-dubbing tasks? How satisfied were you with employing them in fostering your EFL pronunciation of suprasegmental features and learner autonomy? Please, give reasons to justify your answer.