ISSN 1799-2591
Theory and Practice in Language Studies, Vol. 14, No. 9, pp. 2690-2701, September 2024
DOI: https://doi.org/10.17507/tpls.1409.05

Improving the Pronunciation of Voiceless Consonant Sounds in Vietnamese EFL Undergraduates Using the Simplified Verbotonal Approach

Bui Nguyen Nguyet Minh
Ho Chi Minh City Open University, Vietnam

Andrew Lian
Suranaree University of Technology, Thailand

Luu Thi Mai Vy*
Ho Chi Minh City University of Economics and Finance, Vietnam

***Abstract*—In teaching pronunciation, the traditional articulatory approach, commonly used in Vietnamese classrooms, has shown limitations in addressing the phonetic challenges posed by the differences between Vietnamese and English consonant systems. This study investigates the use of an alternative approach, the Simplified Verbotonal Approach (SVA), in improving the pronunciation of voiceless consonants among Vietnamese EFL undergraduates. The SVA, which emphasizes prosodic features through intensive practice with lowpass filtered speech, was hypothesized to aid learners in producing more accurate voiceless consonants. A mixed-methods quasi-experimental design was employed, involving 70 first-year non-English major students. The control group received instruction using standard pronunciation textbooks, while the experimental group utilized an online platform incorporating SVA principles. Pre- and post-tests assessed participants' pronunciation of voiceless consonants in isolation, sentences, and passages. Semi-structured interviews provided qualitative insights into learners' opinions of the SVA. Quantitative results demonstrated significant improvements in the experimental group's pronunciation accuracy, particularly in sentences and passages. Qualitative data revealed positive student feedback on the SVA. These findings suggest that integrating prosodic training through the SVA can significantly enhance the pronunciation of voiceless consonants in Vietnamese learners, offering a viable alternative to traditional articulatory methods in EFL contexts.**

***Index Terms*—verbotonalism, voiceless consonants, pronunciation, prosodic training**

I. Introduction

The development of English communication skills involves teaching pronunciation, lexical resources, grammar, fluency, and accuracy. Of these elements, pronunciation lies at the very heart of building language skills (O'Brien, 2012). Pronunciation includes the articulation and differentiation of phonemes, known as segmentals, and the integration of speech features that form a tonal system, known as suprasegmentals (Pennington & Rogerson-Revell, 2019). Pronunciation instruction aims to assist learners in recognizing and producing the language's sounds, stress patterns, and intonation, which are crucial for message clarity and idea interpretation. Poor pronunciation can hinder effective communication and the comprehension of messages (Kelly, 2000). Approaches to teaching pronunciation trace back to the 17th century, initially based on the phonics-like principle wherein letters or groups of letters systematically represent the sounds of a language. It is argued that focusing on segmental features, namely vowels and consonants, significantly impacts intelligibility, which therefore is emphasized in pronunciation instruction (Wang, 2022). Indeed, research has shown a direct relationship between accurate pronunciation of minimal pairs and an increase in the level of speakers' intelligibility (Collins & Mees, 2013). Techniques such as repetition of individual sounds, drills, and exercises are used to make it easier for learners to remember, gradually understand, and become fluent in articulating sounds. The prevalence of the articulatory approach can be attributed to a basic assumption that individual phonemes serve as the elemental units of spoken language, with words, sentences, paragraphs, and texts constructed upon this phonemic foundation.

Pronunciation teaching has shifted away from native-speaker norms towards a more relaxed and socially just standard to achieve intelligibility as its ultimate goal (Jenkins & Baker, 2015). However, a standard pronunciation level

* Corresponding Author. Email: vyltm@uef.edu.vn

remains essential. Many Vietnamese learners still struggle to achieve even the baseline pronunciation performance. The differences in phonetic structures between Vietnamese and English may lead to a pronounced accent, which often obscures understanding. To be more specific, "the Vietnamese language is different from the English language in that the former is a tone language whereas the latter is intonational. Thus, Vietnamese learners often have trouble with sentence stress and intonation, and so speak English using a relatively flat tone" (Nguyen & Newton, 2021, p. 79). What is more, English differentiates between voiced and voiceless consonants, a distinction crucial for meaning but less emphasized in Vietnamese. Vietnamese learners also tend to mispronounce consonant clusters containing voiceless plosives (Tran & Nguyen, 2022). They have many difficulties in producing English consonants (Bui et al., 2021; Nguyen, 2002; Nguyen, 2021). To make matters worse, in Vietnam, English teachers tend to prioritize teaching grammar and vocabulary for students to prepare for assessments and examinations. Therefore, under these circumstances, it is essential to search for a more effective and practical approach to helping Vietnamese learners improve their pronunciation, especially in producing intelligible voiceless consonants in English.

Furthermore, prior research regarding investigations in improving English consonants among EFL learners has been focused on the implementation of the articulatory approach, which highlights the use of phonemic contrast (Hazan et al., 2005), phonetic transcription (Jantharaviroj, 2019; Harlika et al., 2018), pronunciation drilling techniques (Temirov, 2014; Watanabe & Dinunzio, 2018). For acoustic analysis, Lambacher (1999) utilized electronic visual feedback to facilitate learners in the visualization of their pronunciation and comparison with native patterns. This analysis also included the movements of the articulators. These researchers tend to prioritize the instruction of sound articulation.

Meanwhile, the impacts of prosodic or suprasegmental features in phonetic correction and pronunciation enhancement in the EFL contexts have garnered empirical support (Cai et al., 2021; García, 2018; He, 2014; Ludovic, 2010; Wen, 2019; Yang, 2016; Zhang, 2005). These studies have leveraged the principles of the Verbotonal Approach (VA), a theory of perception (Guberina, 1972), to design instructional activities for learners to strengthen their pronunciation. The foundational mechanism underpinning this theory was implicit prosody training with the integration of kinesthetic cues, which can foster oral fluency in learners. However, in the realm of English consonant developments, scant attention has been devoted to the use of prosodic features, particularly intonation patterns. Therefore, this study aims to fill the void in the literature. Given the positive effects of VA in pronunciation teaching, this study attempted to examine the impact of its simplified version, or the Simplified Verbotonal Approach (SVA) on the production of voiceless consonants among Vietnamese non-English major undergraduates. Hence, the study addressed the following research questions:

1. How effective is the Simplified Verbotonal Approach in enhancing the production of voiceless consonants in English among Vietnamese non-English major undergraduates compared to the articulatory method?
2. What are the opinions of these students on the Simplified Verbotonal Approach?

II. Literature Review

The Verbotonal Approach

The Verbotonal Approach (VA), developed by Petar Guberina in the 1950s, is prominent in speech therapy and designed to enhance speech sounds' perception and production. This approach uses suprasegmental features such as rhythm, intonation, and stress, essential for natural speech production (Guberina, 1972). Unlike traditional phoneme-based approaches, the VA integrates auditory and kinesthetic cues to improve speech perception and articulation (Asp, 2006). One of the notable benefits of the VA is its multisensory learning component, which incorporates body movements and tactile feedback to aid in auditory discrimination and speech production. This is particularly advantageous for individuals with hearing impairments and second language learners, as it promotes the development of accurate and natural speech patterns (Asp, 2006). The effectiveness of the VA has been documented across various populations, demonstrating its utility in both rehabilitative and educational settings (Faulkner, 2009). Its application in language learning and speech therapy underscores its versatility and significant impact on communicative abilities.

A substantial body of literature has explored the application of the principles based on the VA in L2 development. Lian (1980) was among the pioneers who introduced the VA into teaching French pronunciation. He used various techniques such as relaxation exercises, filtered sentence audition, and movement and gesture to reinforce intonation patterns. García (2018) focused on suprasegmental elements like rhythm and intonation before introducing spelling. This method is designed to foster a relaxed and enjoyable learning atmosphere, facilitating learners' auditory and visual associations and ultimately improving their communicative competencies.

He (2014) combined the VA with computer-assisted language learning to improve the pronunciation skills of Chinese learners. This approach entailed employing low-pass filtered speech to highlight prosodic features, encouraging self-study, and utilizing audio materials to augment learning outcomes. Wen (2019) enhanced the pronunciation of English vowels among Chinese learners through the utilization of filtered audio practice for self-study. The results demonstrated significant improvements in vowel perception and pronunciation. Further empirical investigations, such as that conducted by Cai et al. (2021), have delved into optimizing audio-language input signals for Chinese learners of English, employing dichotic listening principles. This study revealed that specific combinations of filtered and unfiltered audio signals could enhance semantic processing and language learning, significantly improving pronunciation skills. Notably, in Vietnam, the study conducted by Luu et al. (2021) implemented principles derived

from the VT to enhance Vietnamese learners' listening comprehension. The study used low-pass audio filters, combining speech and body movement to re-educate learners' auditory perceptions. Given the positive effects of VA in pronunciation teaching, this study attempted to examine the impact of a simplified version, or the Simplified Verbotonal Approach on the production of voiceless consonants among Vietnamese non-English major undergraduates.

The Simplified Verbotonal Approach

The Simplified Verbotonal Approach (SVA), in essence, is another version of the VA. The SVA mainly focuses on raising learners' awareness of the prosodic features, particularly intonations via intensive practice. To draw learners' attention to the intonational patterns and directly stimulate the right brain (Cai et al., 2021), auditory input was modified using lowpass filtering. In other words, lowpass filtered speech is the audio recording that is filtered to reduce detailed information like specific sounds, meaning, and sentence structure, while keeping elements like pitch, amplitude, and rhythm (Perkins et al., 1996). This manipulation helps make the intonational patterns more salient during listening. In this study, the combination of lowpass filtered speech and unfiltered speech is hypothesized to provide learners with sufficient exposure to intonation patterns for making progress in producing English voiceless consonants.

Intonation patterns, in this study, were developed based on two common ones, which are rising-falling intonation and rising intonation (Chun, 2002). The former is typical of simple declarative sentences, commands, and questions that start with a Wh-word, while the latter is characteristic of yes-or-no questions. As defined by Wells (2006), intonation is the melody of speech, which describes the way a speaker's voice fluctuates to communicate both linguistic and pragmatic meanings. Chun (2002) posits that intonation provides additional cues to convey meanings thanks to its multifunctional facets. These functions consist of signaling grammatical structures, disclosing information organization, conveying emotional nuances, and managing conversational dynamics at the discourse level.

Pronunciation Teaching in Vietnam

There are two major trends in the pedagogy of pronunciation: the bottom-up, phoneme-based segmental orientation, and the top-down or suprasegmental orientation (Pennington & Rogerson-Revell, 2019). The segmental approach posits that teaching individual phonemes first will naturally lead to the development of suprasegmental. Conversely, the suprasegmental approach assumes that once prosodic features are established, the segmental discrimination will naturally follow. Proponents of the top-down approach argue that there is a direct link between prosody and meaning in both the production and comprehension of language and that inappropriate use of prosodic patterns is likely to cause more frequent communication breakdowns (Gilbert, 2008; Jackson & O'Brien, 2011). Nonetheless, in the context of teaching pronunciation, the bottom-up approach appears to be more favored by educators. Teachers assert that learners benefit most from explicit phonetic instruction, progressing from form-focused to meaning-focused tasks (Nguyen & Bui, 2021). Additional studies indicate a preference among Vietnamese teachers for the articulatory approach in pronunciation instruction (Nguyen, 2023; Nguyen & Newton, 2020; Tran & Nguyen, 2020).

Difficulties in Articulating English Voiceless Consonant Sounds

Vietnamese and English language do have some shared consonants such as /b, d, k, m, n, f, v, s, z, h, l/ though each language has its distinct consonant sounds (Tang, 2007). McMahon (2002) stressed the importance of distinguishing between voiced and voiceless sounds, which can be physically felt by placing fingers on the larynx. For instance, the vibration felt while sustaining a [zzzzzzz] sound indicates voicing, as opposed to the absence of vibration with a [sssssss] sound, representing voicelessness. Voiceless consonants, such as /p/, /t/, and /k/, are articulated without vocal cord vibration, typically obstructing airflow in speech production. There are nine voiceless consonant sounds in English: /p/ as in "pen", /t/ as in "top", /k/ as in "cat", /f/ as in "fish", /θ/ as in "thing", /s/ as in "sun", /ʃ/ as in "ship", /tʃ/ as in "chat", /h/ as in "hot". These consonants, along with consonant clusters, present challenges for non-native English speakers, including Vietnamese learners. Vietnamese, being a tonal language with fewer voiceless consonants, may not adequately prepare speakers for the articulatory demands of English. Additionally, the lack of phonemic voicing contrast in Vietnamese complicates the perception and production of voiceless sounds in English. Numerous studies have examined the difficulties Vietnamese speakers encounter in pronouncing English consonants and clusters.

Nguyen (2002) conducted a study on Vietnamese L2 learners of English, identifying final consonant clusters that posed challenges in accurate production. The research revealed that clusters containing a liquid (/rt/, /lθ/) were notably more difficult than those with a nasal. Bui's (2016) findings also revealed that the pronunciation of the consonant /θ/ was often substituted with the Vietnamese sound /t'/. Tran (2021) further elucidated that plosive consonants were frequently mispronounced by Vietnamese students. This mispronunciation often involved omitting final sounds such as /t/, /z/, /s/, /k/, and /v/, a habit influenced by the absence of final sounds in Vietnamese pronunciation. Bui et al. (2021) found that Vietnamese sophomores majoring in English frequently erred in pronouncing final consonants, particularly /s/, /z/, /ʃ/, /f/, and /v/, with omission and substitution being the main types of mistakes. Tran and Nguyen (2022) identified consonant clusters containing voiceless plosives as leading to the highest rate of mispronunciation among Vietnamese learners. Moreover, Nguyen and Tran (2023) discovered that stop and fricative consonants, including /b/, /k/, /p/, /t/, /d/, /ʃ/, /v/, and /s/, were commonly mispronounced by a majority of students.

These studies demonstrated that Vietnamese learners do have problematic pronunciation in producing voiceless consonant sounds in English. As Duong (2009) expounded, the confusion surrounding consonant sounds can be attributed to the difficulty in differentiating between sounds, influence from the mother tongue, perception of mistakes, and inadequate drilling and practice.

III. Methodology

This study adopted a mixed-method approach with a quasi-experimental design. Quasi-experimental research often takes place in natural settings without the artificial constraints of a laboratory, providing insights into how the teaching approach functions in real-world conditions (Cook & Campbell, 1979).

Participants

The target population of this study was first-year non-English first-year students. These students were chosen because they represented a broader range of English learners in Vietnam, have less exposure to pronunciation instructions, and tend to have problematic pronunciations, which allows for more precise measurement of improvement for this study. After recruiting 200 students to join a course to improve their pronunciation, piloting was conducted among 100 students. After that, another cohort of 70 students participated in the experiment as a result of convenience sampling. The participants had an average of 13 years of English learning in public schools and no private English education. They were randomly divided into control and experimental groups, with 35 students for each and a balanced gender ratio. All participants consented to the study, and ethical clearance was obtained from the university.

Pedagogical Procedures

In this study, two textbooks, "Ship or Sheep" by Baker (2006) and "Better English Pronunciation" by O'Connor (1998) were used to teach for the control group. Students studied two sections weekly, each lasting 2.5 hours, plus at least one hour of self-study at home.

For the control group, the students were taught how to pronounce vowels, diphthongs, triphthongs, and consonants in English sentences, and to correctly place stress and intonation. The coursebook's content is on vowels, consonants, and prosody. The conventional teaching method followed three main steps: presentation, demonstration, and practice. For the experimental group, students were introduced to an online platform for practice in which the contents were embedded. Students practiced intonation patterns with various sentence types such as statements and questions (Appendix A). The online platform comprised two main components. The first is Moodle (v. 4.2), a widely used Learning Management System that manages access to educational materials. The second component is a delivery application created using the Livresq authoring system (https://livresq.com/en/), which facilitates the development of advanced user interfaces and controls over audiovisual content, specifically audio recordings. These resources were integrated systematically into a website which consisted of 10 computer-based lessons. All these lessons were constructed using the SCORM (Advanced Distributed Learning Initiative) protocol, which means that they can transmit information regarding the completion of exercises and other relevant data to the LMS. Students' access and performance are tracked to ensure compliance with the experiment's directives.

Overall, the procedure consists of eight main steps.

Step 1: When students log in, they will see a list of contents, comprising ten lessons, each featuring various types of sentences: statements, yes-no questions and information questions. Students are required to follow the lessons sequentially and are permitted to revisit previous lessons.

Step 2: When students select the title of a unit, the system automatically presents the content of each lesson, which comprises five sentences. Students must complete sentence 01 before the content of sentence 02 becomes accessible. They are permitted to return to the previous sentence at any time.

Step 3: When students click on Sentence 01, two buttons "Play" and "Reset" will appear. When the students click the PLAY button, they will hear filtered audio of sentence 01 repeated 15 times. This filtering was designed to raise students' awareness of its prosodic characteristics. The students are encouraged to listen to the filtered sentence and, if they wish, synchronize their bodily movements to the sentence's prosody.

Step 4: The recording stops after the students have heard the filtered pattern 15 times. At this point, students must decide whether the audio they heard is one of the displayed options: a yes-no question, an information question, or a statement.

Step 5: After that, students listen to the filtered audio another ten times.

Step 6: Next, students listen to the unfiltered audio to identify the contrast, including similarities and differences, between the filtered and unfiltered versions. They also have the option to display the text of the sentence. Students are encouraged to hum the sentences, either internally or externally.

Step 7: After that, students can listen to both the filtered and unfiltered versions of the sentence, which allows them to review and compare. This step integrates both prosodic and grammatical information, which will help to create perceptual expectations in both the reception and production of natural language, thus enhancing language processing and production. Although the primary focus of the lesson protocol is on enhancing receptive skills - i.e., listening and refining perceptual mechanisms- students can record their voices, compare their recordings with the original model, and download their recordings if they wish.

Research Instruments

There were two main instruments: the voiceless consonant test and semi-structured interviews.

The voiceless consonant test was used as a pre-test and post-test (Appendix B). It consisted of three parts. Part 1 required students to pronounce 27 single words with voiceless consonants in initial, middle, or final positions. Part 2 included reading 104 sentences of varying intonation patterns containing these consonants. Part 3 involved reading a passage highlighting prosodic features like rhythm, stress, and intonation, with the consonants placed in different

positions. Grading Part 1 and 2 (word and sentence level) involved marking word pronunciation as "Correct" or "Incorrect". For grading Part 3 (passage level), criteria from the IELTS speaking test were utilized. The grading scale assessed pronunciation accuracy and intelligibility, ranging from "Intrusive" (0-20%) with severe comprehension hindrance to "Very Good" (80-100%) with nearly flawless pronunciation. Two Vietnamese English teachers with IELTS Speaking Band 8 rated the pre-tests and post-tests independently. The students' recordings were anonymized and randomized and the ratings were entirely blind with the raters not knowing whether they were listening to the experimental or control group, pre-test or post-test. This blind rating technique minimized or eliminated bias because it ensured that the evaluations were based entirely on criteria rather than the knowledge of the participants. The Pearson Correlation Coefficient measuring the consistency between the two raters was calculated, which demonstrated a high value indicating strong agreement and confirming the assessments' reliability and objectivity.

A semi-structured interview was utilized to gather comprehensive data on students' opinions on the implementation of the SVA (Appendix C). Semi-structured interviews allowed interviewees flexibility in questioning and clarity (Ary et al., 2010). Open-ended questions let participants freely express their views on the approach's effectiveness and usefulness, ensuring authentic and detailed responses. Expert evaluation of the interview stages and questions ensured methodological rigor. Interviewing all 35 participants in the experimental group enhanced data reliability and provided a comprehensive understanding of diverse participant experiences, preventing selection bias.

Data Collection and Analysis

For quantitative data, after all the scores were collected, and coded into the SPSS 23, paired sample and independent sample t-tests were conducted to assess differences in participants' mean scores between the pretest and posttest phases. These analyses aimed to identify statistically significant variations in mean scores, thereby evaluating the impact of pronunciation enhancement using the Simplified Verbotonal Approach and the articulatory approach at the group level. The Shapiro-Wilk test results indicated that, except for one case, the data did not follow a normal distribution. Consequently, it was prudent to move away from conventional t-test analysis. The Mann-Whitney U test mitigated the impact of non-normality, offering a more accurate reflection of central tendencies across different instructional groups. ANCOVA was performed in one instance to evaluate precisely the impact of the intervention. For qualitative analysis, after the interviews were carried out, the recordings were transcribed and analyzed using content analysis. Two raters worked on the data separately and then reached the final agreement on the final categorization of ideas based on raw data.

IV. Findings

Quantitative Data

As displayed in Table 1, differences between the means among pretests and posttests regarding each section between the Control Group (C-GRP) and the Experimental Group (E-GRP). The results provide general observations regarding three aspects. First, for improvement patterns, the E-GRP consistently shows larger improvements in mean scores across all tests compared to the C-GRP, suggesting a more effective intervention or different influencing factors. Second, in terms of variability, the standard deviations generally decrease from pre-test to post-test for both groups, indicating more consistent performance post-intervention. However, the E-GRP often ends up with a lower SD, especially in the post-voiceless sounds and voiceless-sound sentences sections. Finally, concerning the performance trends, the E-GRP's notable improvement in voiceless-sound passage post-test scores, compared to the C-GRP, highlights a particularly strong performance in this area.

Table 1
Descriptive Statistics

Test	C-GRP Mean	C-GRP SD	E-GRP Mean	E-GRP SD
Pre-Voiceless sounds	22.26	3.02	21.80	3.71
Post-Voiceless sounds	23.00	2.91	24.04	1.65
Pre-Voiceless-sound Sentences	74.26	13.51	67.13	20.51
Post-Voiceless-sound Sentences	76.69	10.75	79.64	13.29
Pre-Voiceless-sound Passage	62.76	22.42	65.06	24.73
Post-Voiceless-sound Passage	60.56	19.36	72.96	20.94

Note: C-GRP: Control group; E-GRP: experimental group

At the descriptive level, these observations suggest that the E-GRP outperforms the C-GRP in terms of mean score increases and consistent performance across various tests. These results indicate the effectiveness of the intervention used with the E-GRP, which is the effectiveness of SVA in improving voiceless consonant sounds in English.

Data analyses run by the Mann-Whitney U test, as illustrated in Table 2.

TABLE 2
MANN-WHITNEY U TEST (INDEPENDENT SAMPLES TEST)

Test	C-GRP Mean	C-GRP SD	E-GRP Mean	E-GRP SD	p-value	Effect Size
Pre-Voiceless sounds	22.26	3.02	21.80	3.71	0.768	0.042
Post-Voiceless sounds	23.00	2.91	24.04	1.65	0.12	0.22
Pre-Voiceless-sound Sentences	74.26	13.51	67.13	20.51	0.19	0.18
Post-Voiceless-sound Sentences	76.69	10.75	79.64	13.29	0.14 **ANCOVA** <0.001	0.22 **ANCOVA** Cohen's d 0.927
Pre-Voiceless-sound Passage	62.76	22.42	65.06	24.73	0.69	0.34
Post-Voiceless-sound Passage	60.56	19.36	72.96	20.94	0.014	0.62

Note: C-GRP: Control group; E-GRP: experimental group

Voiceless Sounds

In the pre-test of voiceless sounds, there is no significant difference between the control and experimental groups at the pre-test stage ($p > 0.05$). The very small effect size suggests negligible initial differences. In the post-test of voiceless sounds, there is no significant difference between the control and experimental groups at the post-test stage ($p > 0.05$). However, both comparisons of the means C-GRP and E-GRP Post-Voiceless Sounds indicate a small advantage in favour of the experimental group. The small to moderate effect size also indicates some improvement in the experimental group.

Voiceless-Sound Sentences

In the pre-test of voiceless sound sentences, initial calculations suggest that there is no significant difference between the control and experimental groups at the pre-test stage ($p > 0.05$). The small effect size indicates minor initial differences. However, in reality, there is a large difference in the pretest scores between C-GRP (Mean = 74.26, SD = 13.51) and E-GRP (Mean = 67.13, SD = 20.51) in favour of the C-GRP. While in the significance calculations, this difference in Mean of 7.13 was computed as non-significant, this very large difference in favour of the C-GRP seemed to flag an anomaly in the calculations worth investigating.

In the post-test of voiceless sound sentences, calculations suggest that there is no significant difference between the control and experimental groups at the post-test stage ($p > 0.05$). The C-GRP Mean is 76.69 (SD = 10.75) and the E-GRP Mean is 79.64 (SD = 13.29). In light of the initial large difference in the pretest mean for the C-GRP, this translates into the experimental group making up the very large difference and still overtaking the C-GRP by a moderate margin. Also, the moderate effect size suggests improvement in the E-GRP. This turnaround of 10.98 points prompted an ANCOVA analysis to take account of the large pretest difference in scores in assessing the post-test outcome. An ANCOVA analysis for the post-test of voiceless sound sentences was performed and showed a p-value: $p < 0.001$ with effect size (Cohen's d): 0.927. Thus, after adjusting for the pre-test discrepancy using ANCOVA, the results show a significant difference between the control and experimental groups ($p < 0.001$). The large effect size indicates substantial improvement in the experimental group compared to the control group, thus vindicating the statistical concern. In other words, in this particular test, the E-GRP significantly outperformed the C-GRP in the group analysis.

Voiceless-Sound Passage

In the pre-test of sound passage, there is no significant difference between the control and experimental groups at the pre-test stage ($p > 0.05$). In the post-test of sound passage, there is a significant difference between the control and experimental groups as indicated by the difference between the means: 12.40 in favour of E-GRP at the post-test stage ($p=0.014$) and a moderate to a large effect size of 0.62.

In sum, the E-GRP shows significant improvements over the C-GRP in two of the three sections applied with a moderate to large effect size. ANCOVA results for the voiceless-sound sentences indicate a significant difference favoring the E-GRP after adjusting for pre-test discrepancies ($p < 0.001$, Cohen's $d = 0.927$). Regarding the effect size, the rank biserial correlation effect size calculation used standardly in conjunction with the Mann-Whitney U test generally yields lower effect sizes compared to Cohen's d. Nevertheless, moderate to large effect sizes in the post-tests for all three tests indicate substantial improvements in the E-GRP. The effect sizes for pre-tests are generally small, suggesting that initial differences between groups were minor. For the adjustments and robustness, the use of ANCOVA to adjust for pre-test discrepancies in the voiceless-sound sentences test provides a robust and well-established procedure for addressing initial imbalances and confirms the significant improvement in the E-GRP.

The above results indicate that the E-GRP shows greater improvements compared to the C-GRP, particularly in the voiceless-sound passage and voiceless-sound sentences sections. The use of ANCOVA highlights the substantial impact of the intervention on the E-GRP, confirming the effectiveness of the experimental approach in improving voiceless consonant production. The effect sizes calculated using rank biserial correlation are lower than Cohen's d but indicate meaningful differences in outcomes between the two groups.

Qualitative Data

Overall, the analysis of the interview data reveals positive opinions regarding the implementation of the SVA. Participants expressed positive sentiments, citing their interest in the SVA, as well as acknowledging its utility and

efficacy in enhancing their pronunciation of voiceless consonant sounds. All 35 participants expressed excitement about the SVA-based activities, finding the approach innovative and engaging. As illustrated by the remarks of Participant 4, who said, “I quite like this approach and can grasp more effective ways to improve pronunciation." Participant 12 remarked, "It helps me feel that it is not boring during the learning process." Most participants (33 out of 35) acknowledged the approach's usefulness in improving scores and understanding the sounds of the target language. They found it more effective than the articulatory approach. Participants noted measurable pronunciation advancements and appreciated the approach's comprehensive nature as demonstrated by some excerpts from the interviewees:

> It is a very unique course. It is different and fascinating. After studying the course, I really like this method. I can listen to the intonation part and then pronounce it. I can learn to pronounce many words, realize the sounds and pronounce them better than before. This is a very useful approach. (Participant 22)
>
> Learning by this approach will support learners and naturally develop their speaking skills or pronunciation. I can listen to and grasp the intonations of native speakers and the way they emphasize sentences and pronounce linking sounds. Hence, this is a useful way of learning. (Participant 15)

Participants overwhelmingly praised the approach's effectiveness, with 32 out of 35 noting significant improvements. They reported better voice modulation, pronunciation, rhythm, and enhanced listening skills. As some participants articulated:

> This approach will help me adjust my voice better, pronounce better, and have more rhythm when speaking, making my speaking also more rhythmic. Besides, practicing pronunciation like this also helps me listen better and helps me practice listening skills for my upcoming exam preparation. (Participant 23)
>
> After 10 English lessons, I feel that my ability to respond and pronounce vocabulary has improved. It is not about pronouncing each word separately. Moreover, I can combine two new words. My progress is very clear. (Participant 30)

Additionally, Participant 34 observed a significant improvement in articulation, leading to clearer speech: "This approach has made a big difference in how I pronounce the voiceless sounds in English. It is now much easier for me to articulate these sounds accurately, which makes my speech clearer." Participant 19 also reported, "Before taking this course, I struggled with voiceless sounds a lot. Now, I can pronounce them more naturally, which has greatly enhanced my communication ability." She said that a natural ease in producing these sounds greatly enhanced their communication skills.

In brief, the SVA has proven to be a highly effective and engaging approach for enhancing English pronunciation, as evidenced by interviews with 35 students from an experimental cohort. Participants expressed enthusiasm for SVA activities' innovative and interactive nature, noting substantial improvements in pronunciation, fluency, and confidence. They found the approach more effective, appreciating its comprehensive nature and practical benefits.

V. Discussion

Both quantitative and qualitative data indicate the effectiveness of using the SVA in enhancing the production of English voiceless consonant sounds among Vietnamese learners compared to the articulatory approach. These findings lend support to previous research that highlighted the efficacy of the top-down approaches to teaching pronunciation (Gilbert, 2008; Jackson & O’Brien, 2011). The results demonstrate that suprasegmental features should be given precedence in pronunciation instructions since they can facilitate the production of segmental components, particularly voiceless consonants. This is consistent with prior studies (e.g., García, 2018; He, 2018; Lian, 1980; Yang, 2016) that underscored the effectiveness of VA in pronunciation development.

These outcomes suggest the direct link between speech perception and sound articulation which indicates that exerting a certain influence on perception may lead to a change in production. Although the current study did not emphasize kinesthetic elements, raising awareness of prosodic patterns proved to be an effective way to enhance learners’ pronunciation. The utilization of lowpass filtered speech was again shown to be useful for helping L2 learners internalize prosodic patterns (Cai et al., 2021; Luu et al., 2021). In essence, a combination of filtered and unfiltered audio signals may boost semantic processing and language acquisition, leading to noticeable improvements in pronunciation. Salient progress was observed in the performance of voiceless sounds at both sentence and passage levels. Furthermore, the positive feedback from participants in this study also corroborates the opinions of participants in other research (He, 2018; Luu et al., 2021; Yang, 2016), which showed favor for the VA over traditional teaching. Participants appreciated the novelty, value, and efficiency of this approach, resulting in their increased engagement and commitment to fulfilling all required tasks for better outcomes.

These findings have significant implications for pronunciation pedagogy. First, more attention should be given to the use of suprasegmentals in providing pronunciation instructions in educational institutions. Given the prevalence of the articulatory approach in mainstream teaching (Hazan et al., 2005; Harlika et al., 2018; Nguyen & Bui, 2021; Nguyen, 2023), workshops and training should be organized to disseminate this research-based evidence to language educators and learners. Considering the difficulties Vietnamese learners face when producing consonant sounds (Bui et al., 2021; Nguyen & Tran, 2023; Tran, 2021), the principles of VA or SVA should be incorporated into pronunciation practice tasks. Second, the main principles of the VA or SVA should be widely introduced to both researchers and teachers for

further experimentation to validate the effectiveness of this approach. Notably, the use of lowpass filtered speech to raise learners' awareness of prosodic patterns can be applied both within and outside classrooms.

VI. Conclusions

This study investigated the implementation of the SVA in improving the pronunciation of English voiceless consonants among 70 Vietnamese non-English major undergraduates. The study adopted a mixed-method quasi-experimental design. The semi-structured interview was used to gather deeper insights into the use of this approach in pronunciation training. The primary finding that emerged from both quantitative and qualitative data analysis was the effectiveness of the SVA in enhancing voiceless consonant sound production in English compared to the articulatory approach. These empirical findings in this study contribute to our understanding of how pronunciation, particularly voiceless consonants can be developed through extensive exposure to prosodic features. The benefits of this exposure are maximized by using a combination of unfiltered and filtered speech. This work also adds to the growing body of research that indicates the connection between perception and production in language acquisition. Notably, the present study is the first empirical investigation into the impact of SVA on improving pronunciation in Vietnam.

Despite these significant contributions, the study has limitations. Since the participants in this study were recruited in a specific area of Vietnam, the generalizability of these results should be approached with caution for other regions with different learner populations. More research can be done for more validation. Another area of valuable investigation involves the calculation of change scores at an individual level, which would be a fruitful area for future work.

Appendix A Teaching Contents for Two Groups

Learning contents for the control group

Weeks	Contents
1.	Introduction of The Course - Vowels
2.	Vowels, Diphthongs
3.	Triphthongs, English Consonants- Plosive Consonants
4.	Consonants
5.	Nasal Consonants, Lateral Consonants
6.	Consonant Clusters, Stress
7.	Stress (cont.), Connected Speech
8.	Connected Speech (cont.), Intonation
9.	Intonation (cont.)
10.	Review prosody

Learning contents for the experimental group

The systematic study of intonation patterns		Content
1. Short statement	–	Winter is cold
2. Medium length statement	–	The autumn leaves were gently falling from the trees.
3. Long statement	–	In a rapidly changing world, the most robust and intelligent people will survive, as they can best manage change.
4. Long statement containing lists	–	Her morning routine included meditation, a brisk jog, drinking green tea, writing a journal, and reading the news.
5. Short yes/no question	–	Is the sky blue today?
6. Medium-length yes/no question	–	Have you ever considered adopting a vegetarian lifestyle?
7. Short information question	–	Who painted the Mona Lisa?
8. Medium-length information question	–	Which movie won the Best Picture Oscar in 2020?
9. Imperative statement	–	Put on your seatbelt before starting the car
10. Exclamation	–	That is amazing!

APPENDIX B TEST OF VOICELESS CONSONANT SOUNDS

Part I: Word reading
There are 27 words in this part. Please read them out.

/p/	/t/	/k/	/s/	/f/	/θ/	/ʃ/	/ʧ/	/h/
Pot	Ten	Cat	Sun	Fan	Thin	Ship	Chair	Hat
Apple	Water	Rocket	Castle	Office	Athlete	Fashion	Nature	Ahead
Top	Cat	Back	Grass	Leaf	Bath	Bush	Watch	Na

Part II: Sentence reading
Read the following sentences

1.	Fix the flat tire.
2.	Sit right here.
3.	Finish your dish.
4.	The office is off limits.
5.	Did you see the match?
6.	Who is the president of the business?
7.	Is Fiona feeling fine?
8.	Cross the street.
9.	Is it hotter today?
10.	Is this the right bath?
11.	Who lives in the white house?
12.	Where did you find the sock?
13.	The cushion is comfortable.
14.	Ken knows the key.
15.	Peter probably knows.
16.	Who is the official referee?
17.	How high is the hill?
18.	Did Peter play piano?
19.	Who is the cashier in this shop?
20.	Where did you place the keys?
21.	Is Sally at the seaside?
22.	Please use the scissors.
23.	He bathes in the bathtub.
24.	Where is the cat?
25.	Which kite did Karen choose?
26.	Push the button.
27.	Wait a little longer.
28.	Which jacket did Jack pick?
29.	Peter, pass the papers.
30.	Check the package.
31.	Which teacher taught that?
32.	Did you close the suitcase?
33.	Which patch did you pick?
34.	Attach the file.
35.	Which theme did they choose?
36.	Should Shawn shut the door?
37.	Which brush do you use?
38.	Is this your dish?
39.	Did you wash the dishes?
40.	He has a brief.
41.	The locker is locked.
42.	When did Tom talk?
43.	Which chair did Chelsea pick?
44.	Kick the ball.
45.	Turn the music off.
46.	I bought a watch.
47.	Has Henry arrived?
48.	He wears shoes.
49.	Which cloth did you choose?
50.	Tom, tie the knot.
51.	She likes the book.
52.	Hold the handle.
53.	Did Tom take the toy?
54.	Which shoes did Sheila choose?
55.	Walk the path carefully.
56.	Why did Peter panic?
57.	Is the apple ripe?
58.	Is Anthony coming with us?
59.	Did you see the duck?
60.	Hannah has a hobby.
61.	Did Chad choose this?
62.	Frank finds it funny.
63.	The apple pie is delicious.
64.	Wash your hands.
65.	Catch the ball.
66.	Did you refill the coffee?
67.	Offer her a coffee.
68.	Which flavor does Fred favor?
69.	They are behind the house.
70.	Breathe through the straw.
71.	Is that your hat?
72.	Throw the ball.
73.	Is the machine working?
74.	Is this your thing?
75.	Tom tries to travel.
76.	Show me the way.
77.	She should shop more.
78.	Did you watch the match?
79.	Which song did Susan sing?
80.	Is this your cap?
81.	The water bottle is full.
82.	Is the monkey, okay?
83.	Is this leaf green?
84.	The kitchen is clean.
85.	Peel the apples please.
86.	Charlie checks the chart.
87.	Are these your glasses?
88.	She took the path.
89.	Which hotel did you book?
90.	Which apple type did you pick?
91.	The desert has cactuses.
92.	Can Kevin keep it?

93.	Send the message.
94.	Theresa thinks thoroughly.
95.	Check the backpack.
96.	He has a brush.
97.	Place it on the rack.
98.	I left the cap on the table.

99.	Put the cap back.
100.	Where did you place the cap?
101.	Which method did Martha use?
102.	She sat on the mat.
103.	Sam sees the sun.
104.	Where did you find the scarf?

Part III: Passage reading.

Read the following passage.

Peter's cat, named Rocket, loved to lay atop the chair near the window, basking in the sun. One day, as Rocket was napping. Peter decided to repot an apple tree sapling. He took a pot and filled it with water, making the soil moist and ready. On his office desk sat a fan decorated with images of the nature reserve he visited last summer. The fan reminded him of the castle they toured and the ship they boarded. Beside it was a framed picture of his wife, an accomplished athlete, looking as thin and fit as ever. Peter, inspired by a fashion magazine? Yes, he decided to wear his favorite hat for the evening walk. He made a mental note to watch out for the bush near the pond; last time, he ended up taking an unplanned bath. Ahead of him? No, the grass glistened, wet from the recent rain, and the leaves from the tall trees, including a leaf from the oak, blanketed the path.

APPENDIX C QUESTIONS FOR INTERVIEW

Question 1: What are your thoughts on the SVA after engaging with it?

Question 2: Do you believe the SVA is effective in enhancing the pronunciation of voiceless consonants? If so, in what ways?

Question 3: Has the SVA been beneficial for your pronunciation learning? If so, in what ways? If not, why not?

Question 4: How does the SVA compare to the convention method (the articulatory method) in your opinion?

Question 5: Is there anything else you would like to share about the SVA?

REFERENCES

[1] Ary, D., Jacobs, L. C., & Sorensen, C. K. (2010). *Introduction to Research in Education*. Wadsworth.

[2] Asp, C. W. (2006). *Verbotonal speech treatment*. Plural Publishing.

[3] Baker, A. (2006). *Ship or sheep? Book and audio CD Pack: An intermediate pronunciation course*. Cambridge University Press.

[4] Bui, S. T. (2016). Pronunciation of Consonants /ð/and /θ/ by Adult Vietnamese EFL Learners. *Indonesian Journal of Applied Linguistics*, *6*(1), 125-134.

[5] Bui, T. T. L., Mai, T. H., & Diep, H. N. (2021). Common Errors in Pronouncing Final Consonants of English-Majored Sophomores At Tay Do University, Vietnam. *European Journal of English Language Teaching*, *6*(3), 120–161.

[6] Cai, X., Lian, A., Puakpong, N., Shi, Y., Chen, H., Zeng, Y. & Mo, Y. (2021). Optimizing auditory input for foreign language learners through a verbotonal-based dichotic listening approach. *Asian-Pacific Journal of Second and Foreign Language Education*, *6*(2), 1-20.

[7] Carrera-Sabaté, P., Aguadé, C., & Borràs-Comes, J. (2023). Cos i pronúncia: Un tàndem imprescindible? El cos com a eina facilitadora per millorar la pro-núncia del català (Enhancing Speech Production and Perception through Verbotonal Principles). *Revista del Congrés Internacional de Docència Universitària i Innovació (CIDUI)*, *6*(1), Article 418066.

[8] Chun, D. M. (2002). *Discourse Intonation in L2: From Theory and Research to Practice*. John Benjamins Publishing Company.

[9] Cook T. D., & Campbell D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Houghton Mifflin.

[10] Collins, B., & Mees, I. M. (2013). *Practical phonetics and phonology: A resource book for students*. Routledge.

[11] Duong, T. N. (2009). Mistake or Vietnamese English. *VNU Journal of Science*, *25*(4), 41–50.

[12] Denscombe, M. (2017). *The good research guide: For small-scale social research projects*. McGraw-Hill Education.

[13] Dahmen, S., Grice, M., & Roessig, S. (2023). Prosodic and segmental aspects of pronunciation training and their effects on L2. *Languages*, *8*(1), 1-28.

[14] Faulkner, A. (2009). The Verbotonal Approach in Modern Speech Therapy. *International Journal of Speech Language Pathology*, *11*(1), 56-62.

[15] Gilbert, J. B. (2008). *Teaching pronunciation: Using the prosody pyramid*. Cambridge University Press.

[16] Guberina, P. (1972). *Case studies in the use of restricted bands of frequencies in auditory rehabilitation of deaf*. Zagreb.

[17] García, X. P. (2018). Remarks on verbo-tonal phonetics for a communicative context. *Normas*, *8*(1), 259-271.

[18] Harlika, M. S. U., Saifuddin, M., & Fauziyah, N. (2018). Promoting Students' Accuracy in Pronouncing Consonant Sounds by Using English Pronunciation Software. *Journal of Research in Foreign Language Teaching, 1*(2), 14–24.

[19] Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*(3), 360–378.

[20] He, B., Sangarun, P., & Lian, A. (2015). Improving the English pronunciation of Chinese EFL University students through the integration of CALL and verbotonalism. In 17th International CALL Research Conference: *Task Design and CALL* (pp. 276-285). University of Antwerp.

[21] Jackson, C. N., & O'Brien, M. G. (2011). The interaction between prosody and meaning in second language speech production. *Die Unterrichtspraxis/Teaching German*, *44*(1), 1–11.

[22] Jenkins, J., & Baker, W. (2015). *Developments in English as a Lingua Franca*. De Gruyter.

[23] Kelly, G. (2000). *How to Teach Pronunciation.* Pearson Education.
[24] Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *International Journal of Phytoremediation*, *21*(1), 137–156.
[25] Lian, A. P. (1980). *Intonation patterns of French*. River Seine.
[26] Lian, A. P., & Sangarun, P. (2023). Rhizomatic learning systems and precision language education: A perfect match. In *New Directions in Rhizomatic Learning* (pp. 119-141). Routledge.
[27] Ludovic, K. (2010). Phonetic Correction in Class with Verbo-tonal Method. *Studies in Language and Literature*, *30*(1), 35-56.
[28] Luu, T. M. V., Lian, A. P., & Siriyothin, P. (2021). Developing EFL learners' listening comprehension through a computer-assisted self-regulated prosody-based listening platform. *Computer-Assisted Language Learning Electronic Journal*, *22*(1), 254-271.
[29] McMahon, A. M. (2002). *An Introduction to English Phonology*. Edinburgh University Press.
[30] Nguyen, D. A. D. (2018). Learners' perspectives on English pronunciation teaching and learning: A preliminary study in the Vietnamese context. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*. Iowa State University.
[31] Nguyen, L. T., & Newton, J. (2021). Enhancing EFL Teachers' Pronunciation Pedagogy Through Professional Learning: A Vietnamese Case Study. *RELC Journal, 52(1*), 77-93.
[32] Nguyen, T. P., & Truong, V. T. (2021). Foreign Accent and Speech Intelligibility-Why and How to Reduce. *Asian Journal of Multidisciplinary Research & Review*, *2*(6), 85-102.
[33] Nguyen, H. N. (2023). Teaching English sounds to Vietnamese secondary school students: From theories to applications using learner-centered methods. *International Journal of Language Instruction*, *2*(1), 16-29.
[34] Nguyen, N. (2002). *Interlanguage Phonology and the Pronunciation of English Final Consonant Clusters by Native Speakers of Vietnamese* [Unpublished master thesis]. Ohio University.
[35] Nguyen, T. H. (2021). Contrastive Analysis of Consonants in English and Vietnamese. *Journal of English Language Teaching and Applied Linguistics*, *3*(6), 58–65.
[36] Nguyen, L. T., & Hung, B. P. (2021). Communicative pronunciation teaching: Insights from the Vietnamese tertiary EFL classroom. *System*, *101*, 102537.
[37] O'Brien, M. G. (2012). Pronunciation Matters. *Teaching German*, *37*(1), 1–9.
[38] O'Connor, J. D. (1980). *Better English Pronunciation*. Cambridge University Press.
[39] Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: Contemporary perspectives*. Palgrave Macmillan.
[40] Perkins, J. M., Baran, J. A., & Gandour, J. (1996). Hemispheric specialization in processing intonation contours. *Aphasiology*, *10*(4), 343–36.
[41] Saito, K. (2011). Identifying problematic segmental features to acquire comprehensible pronunciation in EFL settings: the case of Japanese learners of English. *RELC journal*, *42*(3), 363–378.
[42] Tang, G. M. (2007). Cross-Linguistic Analysis of Vietnamese and English with Implications for Vietnamese Language Acquisition and Maintenance in the United States. *Journal of Southeast Asian American Education and Advancement*, *2*(3), 1-33.
[43] Temirov, M. (2014). Improving intelligibility of voiced and voiceless "TH" consonants in the speeches of sophomores using pronunciation drills. *Jurnal Linguistik Terapan Politeknik Negeri Malang*, *4*(2), 53–54.
[44] Tran, D. P. T., & Nguyen, H. B. (2020). EFL teachers' beliefs and practices of teaching pronunciation in a Vietnamese setting. *Universal Journal of Educational Research*, *8*(12), 7022-7035.
[45] Tran, K. M. A. (2021). An Investigation into The Common Pronunciation Errors Made by Young Learners. *Journal of Science and Technology*, *52*(4), 1-11.
[46] Tran, T. K. L., & Nguyen, A. T. (2022). Common Mistakes in Pronouncing English Consonant Clusters: A Case Study of Vietnamese Learners: A Case Study of Vietnamese Learners. *Can Tho University Journal of Science*, *14*(3), 32-39.
[47] Wang, X. (2022). Segmental versus suprasegmental: Which one is more important to teach? *RELC Journal*, *53*(1), 194-202.
[48] Watanabe, K., & Dinunzio, N. (2018). Effectiveness of Teaching Onset Consonant Clusters for Japanese University Learners of English. *International Journal of Curriculum Development and Practice*, *20*(1), 11–19.
[49] Wells, J. C. (2006). *English Intonation: An Introduction*. Cambridge University Press.
[50] Wen, F. (2019). *A verbotonal-based approach to phonetic correction of a selection of English vowels for Chinese EFL learners* [Doctoral dissertation]. Suranaree University of Technology.
[51] Yang, Y. (2016). *Improving the phonological working memory and English-speaking skills of Chinese primary EFL learners with a verbotonal-based approach* [Doctoral dissertation]. Suranaree University of Technology.
[52] Zhang, F. (2005). Using an interactive feedback tool to enhance pronunciation in language learning. In S. Mishra & R. C. Sharma (Eds.), *Interactive multimedia in education and training* (pp. 377–399). Idea Group Publishing.

Bui Nguyen Nguyet Minh is a lecturer at Saigon University, Vietnam. She is also a visiting lecturer at Ho Chi Minh City Open University. Her academic interests encompass pronunciation, translation, and theories in language learning and teaching. Email: bnnminh@sgu.edu.vn

Andrew Lian is a Professor of Foreign Language Studies, the School of Foreign Languages, Suranaree University of Technology, Thailand. He is also Professor Emeritus of Languages and Second Language Education at the University of Canberra, Canberra, Australia. He is the current President of AsiaCALL, a research and professional association focusing on the uses of technology to enhance second/foreign language learning in Asian contexts. His current research interests include neuroscience, perception and

cerebral lateralization as they relate to language learning as well as self-adjusting and self-organizing learning systems based on rhizomatic principles. Email: andrew.lian@andrewlian.com

Luu Thi Mai Vy is a lecturer at Ho Chi Minh City University of Economics and Finance, Vietnam. Her research interests include L2 listening development, pronunciation, and theories in language learning and teaching. Email: vyltm@uef.edu.vn