

# A Review of Pronunciation Instruction in Vietnam: Potential Applications of the Verbotonal Method

*Bui Nguyen Nguyet Minh\**, *Andrew Lian\*\**, *Luu Thi Mai Vy\*\*\**

*\*Saigon University, Vietnam*

*\*\*Suranaree University of Technology, Thailand*

*\*\*\*Ho Chi Minh City University of Economics and Finance, Vietnam*

*Received: 08/7/2024; Accepted: 18/7/2024; Published: 25/7/2024*

**Abstract:** *The challenge of varying phonological systems across languages presents significant difficulties for language learners in achieving the target language. Despite its critical role in effective communication, pronunciation often receives less attention in language instruction than other language skills. This paper highlights the importance of pronunciation instruction and delineates the current obstacles faced in teaching pronunciation within the Vietnamese educational context. It also provides an overview of the Verbotonal Approach (VA), a perception-based theory applied to pronunciation pedagogy. Given that the VA principles reflect the natural processes of language learning, this approach is advocated for teaching pronunciation to Vietnamese learners. The paper concludes with suggestions for developing a protocol for pronunciation practice within a virtual learning environment based on the VA.*

**Keywords:** *Pronunciation, Verbotonal Approach, Review, Prosody, Vietnamese EFL learners*

## 1. Introduction

English is a global means of communication used across various fields, even non-English-speaking countries. Jenkins (2006) noted that “the English language has become the world’s principal lingua franca” (p. 157). Non-native speakers constitute about 80% of English communication (Beneke, 1991). Consequently, English forms vary due to cultural, historical, and regional influences, especially affecting speaking skills. Effective teaching of English speaking involves guiding learners in pronunciation, vocabulary, grammar, fluency, and accuracy. Poor pronunciation hinders comprehension (Kelly, 2000), requiring learners to master segmental and suprasegmental features for intelligibility (Fraser, 2001; Pennington & Rogerson-Revell, 2019).

## 2. Aspects of English Pronunciation

### Segmental Features

Wang (2022) categorises segmental features as individual sounds like vowels and consonants, which are the building blocks of spoken language. These phonemes combine to form words and sentences. Hewings (2004) states that “the building blocks of pronunciation are the individual sounds,” the smallest units in the pronunciation system, such as vowels (/e/, /ə/) and consonants (/tʃ/, /f/). Practising segmental features involves producing accurate speech sounds, understanding phonetic properties, and distinguishing

appropriate sounds in English through phonetic drills, visual displays, and minimal pairs. Researchers like Derwing et al. (1998), Saito and Lyster (2012), and Collins and Mees (2013) emphasise teaching segmentals to improve communication intelligibility.

### Suprasegmental Features

Suprasegmentals, or prosody, encompass features beyond individual speech sounds, such as stress, intonation, pitch, rhythm, loudness, and tempo. Seidlhofer (2001) notes that prosody spans entire utterances, enriching speech’s meaning and emotional expression. Fujisaki (1997) explains that prosody organises linguistic units to convey various types of information. Prosody adds depth, nuance, and expressiveness to communication, allowing speakers to convey meanings beyond literal words through pitch, stress, and intonation. Dahmen et al. (2023) highlight that focusing on prosody improves both prosodic and segmental pronunciation features. Studies by Hahn (2004) and Kang et al. (2010) show suprasegmentals’ importance for intelligibility, making pronunciation instruction more effective. Integrating prosody in teaching enhances clarity and expressiveness.

## 3. Significance of Pronunciation

Pronunciation is crucial for effective communication, ensuring speakers are understood and minimising misunderstandings. Fraser (2001)

emphasises that even minor pronunciation errors can create significant communication barriers, highlighting the need for clear speech. Accurate pronunciation boosts confidence, reduces anxiety, and encourages active discussion participation (Morley, 1994). It also enhances the naturalness and fluency of speech, making communication more engaging and persuasive (Khaleel et al., 2024). Effective pronunciation instruction improves self-esteem and communicative ability, leading to greater language use. Derwing and Munro (2015) argue that the goal should be intelligibility and effective communication, rather than native-like pronunciation, to help learners communicate successfully.

#### **4. Challenges in Teaching Pronunciation in the Vietnamese Context**

Current teaching practices in Vietnam blend conventional and modern methodologies, adapting to global educational trends while addressing local needs. This shift from rote memorisation to dynamic learning aims to improve students' communicative competence in English, which is essential in today's globalised world.

Contemporary Vietnamese classrooms significantly incorporate technology, using digital tools and online resources to supplement traditional textbooks (Gruba & Chau, 2019). Teachers use multimedia presentations and interactive software to engage students, creating a more interactive learning environment. Project-based learning is gaining popularity, allowing students to work on real-world projects that promote critical thinking and problem-solving skills (Tran & Phan, 2021).

Despite these advancements, teaching pronunciation remains challenging. Large class sizes limit individual attention, making it difficult to address specific pronunciation issues. The traditional focus on grammar and vocabulary over oral skills has left some teachers less prepared for effective pronunciation teaching strategies, highlighting the need for ongoing professional development.

Most Vietnamese teachers focus on listening, reading, writing, and speaking skills, with pronunciation as secondary. This perspective necessitates greater emphasis on pronunciation in teaching practices. Studies on pronunciation for Vietnamese learners have largely been survey-based. Phuong (2019) used questionnaires and interviews to investigate Vietnamese learners' views and needs regarding learning pronunciation, recommending

that teachers understand their students' motivations to establish suitable teaching techniques.

Phuong (2022) explored how teachers' feedback on pronunciation affects learners, finding that immediate feedback is prioritised over delayed feedback. Khau and Huynh (2022) gathered information on students' and teachers' opinions on oral fluency, offering suggestions to support correct pronunciation and efficient teaching.

Most studies in Vietnam have focused on segmentals. Tran and Nguyen (2020) noted that teachers mainly teach segmentals, especially problematic sounds like /θ/ and /ð/, often neglecting prosody. Nguyen and Newton (2021) observed that teachers predominantly use listen-and-repeat exercises, emphasising segmentals (Phuong, 2022).

#### **5. An Overview of the Verbotonal Approach in English Teaching**

The Verbotonal Approach, developed by Petar Guberina (1952), emphasises the primacy of spoken language and oral discourse in linguistics and language learning. Guberina posits that language evolves through oral interaction and cannot be separated from its spoken form, aligning with Ferdinand de Saussure's distinction between "langue" (the system of language rules) and "parole" (individual language use). According to Guberina, spoken language reflects the socio-cultural context of its users, including traditions, customs, and societal norms, and is shaped by historical events, cultural exchanges, and social dynamics. When learning a foreign language, individuals often use their native language (L1) to understand the new language (L2), drawing on familiar sounds and patterns. However, this can lead to interference and pronunciation errors in L2.

Guberina (1952) emphasises the importance of non-lexical elements in oral discourse, such as intonation, rhythm, intensity, pitch, pauses, and speech tempo, which add depth and nuance to spoken interactions. By focusing on these non-verbal cues, he highlights the multidimensional nature of spoken language, offering a richer mode of expression than written language. He also stresses the inseparability of verbal speech and body language, arguing that verbal and non-verbal cues form a comprehensive system of global body language used during interactions. This includes physical gestures, facial expressions, and spatial dynamics, which are integral to human communication.

Furthermore, Guberina's framework includes components like Body, Effort, Shape, and Space, applied in dance, theatre, and physical therapy to enhance performance and communication. The Verbotonal Approach thus provides a holistic understanding of dialogue, where meaning is conveyed through a blend of spoken words and physical movements.

### ***Main Principles of The Verbotonal Approach (VA)***

The Verbotonal Approach (VA), grounded in treating children's hearing impairments, remains vital in modern language teaching. Influenced by Trubetzkoy's (1939) linguistic philosophy, VA recognises the phonological sieve, where learners filter unfamiliar sounds through their native language, leading to challenges with new phonetic elements. Ellis (2004) describes language transfer as a cognitive process where learners use L1 knowledge to aid in L2 acquisition, making phonetic perception crucial in VA. Accurate speech mirrors our perception of sounds, with misperception causing mispronunciation (Guberina & Asp, 1981).

VA emphasises suprasegmental elements—pitch, stress, intonation, and rhythm (Lass, 2012). Unlike phonemes, these features are intrinsic to speech and taught first in VA, reflecting infant language acquisition where rhythm and intonation precede phonemic elements. Babies absorb sounds through muscle activity before recognising phonological and prosodic features in speech (Guberina & Asp, 1981).

VA also integrates non-verbal language. Gestures often accompany verbal communication, synchronising naturally with speech to enhance linguistic expression. This natural synchronisation is evident in children whose gestures enhance language production.

The Verbotonal Approach combines phonetic perception, suprasegmental features, and non-verbal language to improve pronunciation. By focusing on sound perception and production, prosody and intonation, and the role of gestures, VA offers a holistic approach for teaching pronunciation, addressing learners' comprehensive communication needs.

### ***Previous Research Related to the Verbotonal Approach***

The Verbotonal Approach (VA) has proven effective in enhancing language learners' communication abilities. Lian (1980) developed a protocol using VA to teach French intonation patterns involving

relaxation, filtered sentence audition, movement and gesture, humming, intonation interpretation, mouthing words, repetition with filtered patterns, and repetition exercises. This method enhanced students' speaking abilities through structured and sensory-rich activities.

He (2014) conducted an experimental study combining VA with computer-assisted language learning to improve Chinese English majors' pronunciation and self-study skills. The study promoted student-centred learning by using filtered audio to highlight intonation and rhythm. Results showed that the experimental group excelled in phonemes, intonation, and stress.

Wen (2019) followed up with a study on enhancing the pronunciation of six specific English vowels among Chinese students. This study emphasised self-study with filtered audio, revealing significant improvements in vowel pronunciation and perception in the experimental group. Each student had unique frequency preferences, highlighting the importance of personalised learning approaches.

Cai et al. (2021) explored optimal audio-language input for Chinese learners using dichotic listening. They tested various combinations by filtering voice signals in one ear and leaving them unfiltered in the other. The FL-R (filtered left/unfiltered right) configuration induced more active semantic processing, enhancing language understanding. Participants using FL-R outperformed others in pronunciation, demonstrating the importance of optimal audio configurations in language learning.

Luu et al. (2021) assessed VA's effectiveness in improving L2 listening skills in Vietnam. The study involved 65 first-year non-English primary Vietnamese university students using low-pass audio filters and integrating speech with body movement. A dedicated website facilitated convenient listening practice, emphasising self-monitoring and self-assessment. After ten weeks, the experimental group showed significantly better listening skills than the control group.

In summary, research into VA application in L2 teaching highlights its potential for enhancing language learning. These studies demonstrate VA's effectiveness in improving pronunciation, speaking, and listening skills, making it a valuable resource for language educators.

### **6. Using the Verbotonal Approach for Pronunciation Development in English**

Given technological advancements and evolving insights into language acquisition, assessing and adapting the Verbotonal Approach (VA) to contemporary needs is crucial. The VA views speech as a comprehensive embodiment involving the entire body, integrating body movement, muscular tension, timing, loudness, rhythm, and spatial awareness to enhance speech potential (Carrera-Sabaté et al., 2023). These elements are vital for improving speech production and auditory perception. Researchers are now focusing on sound filtering techniques, such as using a low-pass filter with a 320Hz cutoff to refine auditory input. This technique clarifies sound by filtering out higher frequencies, making intonation patterns more prominent and stimulating the right brain.

Lian and Sangarun (2023) highlight the demand for rapid, personalised learning facilitated by the Internet and platforms like Google and YouTube. This aligns with rhizomatic learning systems, which stress that learning is inherently personal and varies among individuals. Schools often struggle to meet each student's unique learning needs, making effective customisation challenging. Emphasising a student's "pedagogic need" is crucial, as genuine personalisation requires understanding when a student needs help—a feeling only the student can identify.

In conclusion, the Verbotonal Approach (VA) can enhance pronunciation teaching through key principles and targeted activities that improve learners' awareness and proficiency in English pronunciation.

Firstly, prosody emphasis is crucial, with activities highlighting rhythm, stress patterns, and intonation, thereby improving overall speech coherence and individual sound pronunciation.

Secondly, sound filtering at a 320Hz threshold enhances auditory clarity, allowing learners to focus on fundamental frequencies and intonational contours, reducing auditory complexity and familiarising learners with English prosody.

Thirdly, digital adaptation employs virtual environments and interactive exercises, enabling learners to record and compare their speech with original audio for self-reflection.

Lastly, holistic speech involvement encourages full-body engagement during speech production and auditory perception, promoting the synchronisation

of bodily movements with auditory input through activities like humming and speech audio.

## 7. Conclusion

Pronunciation goes beyond correct sound articulation; it is a crucial communication element that should be embedded in classroom activities. The Verbotonal Approach (VA) offers a comprehensive approach to enhance pronunciation by focusing on musicality and prosody. Innovations in sound filtering and digital VA adaptations underscore pronunciation's significance in language learning. This approach improves understanding of native speakers and overall communication. Given the VA's proven benefits, more empirical research is needed, especially for Vietnamese learners, to refine pronunciation teaching approach tailored to their linguistic and learning needs

## References

1. Beneke, J. (1991). *Englisch als lingua franca oder als Medium interkultureller Kommunikation*. Cornelsen & Oxford University Press.
2. Carrera-Sabaté, J. & Lopez-Garcia, P. & Cols-Aguadé, C. & Borràs-Comes, J. (2023). Cos i pronúncia: Un tàndem imprescindible? El cos com a eina facilitadora per millorar la pronúncia del català. *Journal of the International Congress of University Teaching and Innovation (CIDUI)*, 66.
3. Collins, B., & Mees, I. M. (2013). *Practical phonetics and phonology: A resource book for students*. Routledge.
4. Cai, X., Lian, A., Puakpong, N., Shi, Y., Chen, H., Zeng, Y., & Mo, Y. (2021). Optimising auditory input for foreign language learners through a verbotonal-based dichotic listening approach. *Asian-Pacific Journal of Second and Foreign Language Education*, 6, 1-20.
5. Dahmen, S., Grice, M., & Roessig, S. (2023). Prosodic and Segmental Aspects of Pronunciation Training and Their Effects on L2. *Languages*, 8(1), 74.
6. Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language learning*, 48(3), 393-410.
7. Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.