

Problems in the Pronunciation of English Consonant Clusters – A Case Study of English-Majored Students at a Vietnamese University

Phí Thị Thu Trang

Vietnam Women's Academy

ABSTRACT: Pronunciation plays a decisive role in successful communication, particularly for learners of English as a foreign language. For Vietnamese students, the structural differences between Vietnamese and English phonological systems create significant challenges in mastering English pronunciation. This qualitative study investigates the difficulties faced by English language majors at a Vietnamese university in pronouncing English consonant clusters. Through participatory observation, analysis of recurrent errors and focus group interview, three main tendencies to ease cluster pronunciation were identified: consonant deletion, sound substitution, and vowel insertion. From this, some pedagogical solutions are proposed to address these issues, with the aim of improving students' fluency, accuracy, and confidence in English communication.

KEYWORDS-English pronunciation, consonant clusters, difficulties, English-majored students, coping strategies

I. INTRODUCTION

The importance of teaching and learning English has grown significantly in today's globalized context, as English functions as a lingua-franca across many domains ranging from business to science and as a requirement for individuals to work in a dynamic, ever-changing international environment. Proficiency in English not only facilitates smoother communication but also provides learners with broader academic and professional opportunities. However, mastering English is not without challenges for non-native speakers, particularly in the areas of phonetics and phonology, where the complexity of the sound system often affects both intelligibility and fluency. Learners must devote considerable time and effort to improving their pronunciation in order to achieve comprehensibility and communicative success (García, 2007). For this reason, pronunciation training should be a central goal of language instruction, with teachers guiding learners toward producing sounds as close as possible to those of native speakers. Identifying and analyzing the difficulties learners face in acquiring accurate or at least acceptable pronunciation is therefore an essential task.

As a teacher who worked at a Vietnamese university for over 15 years and taught pronunciation to thousands of English language majors, I observed that students with good pronunciation often achieve much higher overall results in English. I also noticed many types of pronunciation errors, but one of the greatest challenges that most students have not yet overcome is pronouncing consonant clusters in English, a feature that does not exist in Vietnamese. Because Vietnamese and English are typologically different language systems, especially in their phonological systems, Vietnamese learners in general and English-language majors in particular face many difficulties in learning English pronunciation, particularly consonant clusters. Under the interference of their mother language, students tend to substitute English phonemes with Vietnamese ones or drop difficult sounds while speaking, leading to non-standard and even incorrect pronunciation, which affects the transmission of meaning in communication with foreigners. Cook (1991) explains that such adaptations occur because learners attempt to make English conform to the syllable structures of their first language. He emphasizes that pronunciation errors arise more from differences in syllable structure and permissible consonant-vowel combinations than from contrasts between individual phonemes. Apart from this, pronunciation difficulties among foreign students stem from factors such as limited motivation, inadequate exposure to the target language, and shortcomings in instructional practices Gilakjani (2011).

With the goal of improving pronunciation and strengthening speaking abilities among English-major university students, I carried out surveys, classroom observations, and experiential notes to identify the difficulties they faced in producing English consonant clusters. The purpose of this study is to highlight the frequent errors in cluster pronunciation and to suggest practical strategies for achieving clearer and more fluent speech. By doing so, the research seeks to enhance learners' communicative confidence and support the development of other essential language skills.

II. THEORETICAL BACKGROUND

A. Overview of Consonants in English and Vietnamese

Consonants are defined as speech sounds produced with a complete or partial constriction of the vocal tract that modifies the airflow. Their classification is traditionally based on three articulatory parameters: place of articulation, manner of articulation, and voicing. In terms of place of articulation, consonants may be produced at various points along the vocal tract, including bilabial, labiodental, dental, alveolar, post-alveolar, retroflex, palatal, velar, uvular, pharyngeal, and glottal positions. With respect to manner of articulation, they can be categorized as plosives (stops), nasals, trills, taps, fricatives, lateral fricatives, approximants, and lateral approximants. Finally, consonants are distinguished by voicing, depending on whether or not the vocal folds vibrate during production, yielding a contrast between voiced and voiceless sounds.

English has 24 consonants, shown in the table below.

Table 1: English consonant phonemes

		PLACE OF ARTICULATION							
		Bilabial	Labiodental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
MANNER OF ARTICULATION	Plosive	p b			t d				k g
	Fricative		f v	θ ð	s z	ʃ ʒ			h
	Affricate					tʃ dʒ			
	Nasal	m			n			ŋ	
	Lateral approximant				l				
	Approximant	w				r	j		

(Roach, 2009: 52)

In English, consonants are articulated by creating either a partial or complete obstruction of the airstream as it passes from the lungs through the vocal tract (Finegan, 1994). As shown in the table, the majority of English consonants fall into the categories of fricatives, plosives, and nasals, with the addition of the affricates /tʃ/ and /dʒ/, which represent a combination of a stop followed by a fricative. In the chart, each consonant pair is organized such that the voiceless member appears on the left and the corresponding voiced member on the right.

Vietnamese has 21 consonants (including the aspirated /tʰ/, which is an alveolar stop), shown in the annotated table below:

Table 2: Vietnamese consonant phonemes

		PLACE OF ARTICULATION								
		Bilabial	Labiodental	Dental	Denti- alveolar	Post- alveolar	Retroflex	Palatal	Velar	Glottal
MANNER OF ARTICULATION	Plosive	p ɓ			t ɗ tʰ		ʈ	c	k kʰ	
	Fricative		f v		s z		ʂ ʐ		χ ʁ	h
	Affricate									
	Trill				r					
	Nasal	m			n			ɲ	ŋ	
	Approximant	w			l			j		

Problems in the Pronunciation of English Consonant Clusters – A Case Study of English-Majored Students at a Vietnamese University

Articulatorily, most Vietnamese syllables begin with a closing gesture, creating partial or complete obstruction of airflow that yields a particular acoustic effect. Thus, the general property of onsets in Vietnamese syllable structure is their consonantal nature (the bilabial /p/ and the trill /r/ are considered outside the core system, although they exist in certain words or are pronounced in certain regions) (Đoàn, 1999). Comparing the Vietnamese and English consonant inventories, we find some English sounds do not exist in Vietnamese, such as the dental fricatives /θ/ /ð/, post-alveolar sounds, the palatal glide /j/, and the affricates /tʃ/, /dʒ/. Therefore, in theory, Vietnamese learners can be expected to have difficulty pronouncing these consonants, leading to consonant-pronunciation errors that affect communication, especially with final consonants.

B. Consonants clusters in English

In English, many words have more than one consonant combined with a vowel to form a syllable. When two or more consonants are pronounced together within the same syllable, they are called consonant clusters (Roach, 2009). According to Avery and Ehrlich (1992), clusters occur in onset (e.g., play /pl-/ , street /str-/) and coda positions (e.g., texts /-ksts/), with onsets typically limited to two or three segments whereas codas may contain up to four. Initial clusters may have two or three consonants, as in play, cream, street, splash, etc. Final clusters may have from two to four consonants (in cases where the grammatical morphemes -s or -ed are added), as in end, glimpse, twelfths, asked, etc. In connected speech, consonants may also be blended, creating clusters in the middle of a word or between words pronounced together. For example, in pamphlet, the first syllable ends in /m/, and the second syllable begins with /fl/, so when pamphlet is pronounced, /mfl/ occurs medially. Similarly, when pronouncing best boy, the sequence /stb/ appears between the two words, no different from /mfl/ in the two-syllable word pamphlet. Thus, English clusters occur in three basic positions: initial, final, and medial (between two syllables or two words).

Cluster well-formedness is shaped by sonority sequencing and minimal sonority distance: rising sonority in onsets and falling in codas favor learnability, while structures that violate or approach the limits of these patterns (e.g., /s/+stop onsets, dense codas like /-ksts/) are comparatively difficult (Eckman, 1977; Vennemann, 2012).

III. REVIEW OF RELATED RESEARCH

Pronunciation in second language learning has long been a focus of inquiry, with a substantial body of research highlighting how first language (L1) phonological systems shape second language (L2) production (Avery & Ehrlich, 1992; Yavaş, 2005). One of the most persistent difficulties across learner groups involves the production of consonant clusters, which are structurally complex and often absent in learners' native languages (Vennemann, 2012). Therefore, learners adopt a variety of repair strategies, such as deletion, substitution, or vowel epenthesis, to simplify clusters, resulting in systematic pronunciation errors (Anderson, 1987; Major, 1996).

Cross-linguistic studies confirm that cluster difficulty varies by L1 background. For example, Korean learners, whose language does not permit onset clusters, frequently insert vowels in /sC/ onsets, producing forms such as /əsport/ for sport (Choi, 2016). Similarly, Arabic learners, whose phonology restricts final clusters, often employ epenthesis or deletion to accommodate English coda structures (Altakhaineh, Al-Junaid, & Younes, 2024). Persian learners have also been found to struggle with onset clusters, showing high rates of vowel insertion in words like spring or train, reflecting the absence of initial clusters in Persian (Khanbeiki & Abdolmanafi-Rokni, 2015). Yemeni learners in Malaysia demonstrated similar patterns but improved significantly after targeted training, indicating that instruction can mitigate L1 transfer effects (Baagbah, Jaganathan, & Mohamad, 2019).

Research on Vietnamese learners echoes these findings while highlighting unique difficulties. Vietnamese does not permit consonant clusters in onset position and allows only a limited set in coda position, which means English clusters are highly marked for these learners (Sato, 1984; Benson, 1986). Empirical studies have reported that Vietnamese learners often delete initial consonants in clusters or alter voicing features of plosives, with epenthesis occurring less frequently than in other L1 groups (Tran & Nguyen, 2022). Ngo (2011) and Nguyen (2020) noted that learners consistently omit or reduce final clusters such as /-st/ or /-nd/, demonstrating the influence of Vietnamese phonotactics. These findings suggest that L1 transfer is the primary source of error, in line with Eckman's (1977) Markedness Differential Hypothesis, which predicts that more marked structures absent in the L1 will be especially difficult for L2 learners.

In conclusion, the literature shows that consonant cluster pronunciation is a cross-linguistic challenge, but its manifestation depends on specific L1 constraints. For Vietnamese learners in particular, the absence of onset clusters and the limited nature of final codas lead to predictable errors, most notably deletion and feature change. While some studies have described these tendencies, systematic investigations focusing on English majors in Vietnam remain limited. The present study builds on previous work by providing a detailed case study of English-majored students at a Vietnamese university, aiming to document their common errors with clusters and identify pedagogical strategies to address them.

Problems in the Pronunciation of English Consonant Clusters – A Case Study of English-Majored Students at a Vietnamese University

IV. PARTICIPANTS AND METHODS

The study involved 162 first-year English majors at a university in northern Vietnam, carried out during their second compulsory pronunciation course as part of the official curriculum. Across two pronunciation courses, students were introduced to fundamental concepts including letters, sounds, syllables, word stress, and intonation, and they practiced these features under the teacher's guidance in class. Table 3 below summarizes background information about the participants.

Table 3: Participant information

Category	n
English learning experience	
< 5 years	3
5–10 years	64
≥ 10 years	95
Pronunciation course before university	
Have experience	63
No experience	99

In terms of English learning experience, most students reported studying the language for at least ten years ($n = 95$), followed by 64 students with 5-10 years of experience, and only 3 with fewer than five years. Regarding prior pronunciation training, 63 participants had taken a pronunciation course before entering university, whereas 99 reported no such experience. This distribution highlights a participant pool with substantial English learning backgrounds but limited formal exposure to pronunciation-focused instruction.

The study was carried out by the researcher, who also served as the instructor responsible for teaching pronunciation to the participants. A qualitative research design was adopted to explore learners' challenges in producing English consonant clusters. Data were gathered through classroom observations, systematic analysis of recurrent pronunciation errors, and in-depth interviews with selected students who experienced particular difficulty with cluster production. These combined methods facilitated the identification of prevalent error patterns and learners' coping strategies, thereby providing a foundation for proposing pedagogical interventions aimed at improving students' pronunciation accuracy and fluency. The purpose of the study was not to generate quantitative data but rather to highlight insights of pedagogical significance. Its intention was to provide English language teachers with thought-provoking perspectives that could guide the development of effective instructional strategies aimed at enhancing learners' pronunciation as well as their broader language competencies.

V. FINDINGS AND DISCUSSION

From a theoretical perspective, there are three key factors that contribute to students' difficulties with consonant clusters. First, differences between the English and Vietnamese consonant inventories create challenges: certain English consonants are absent in Vietnamese, while others exist but differ in their phonetic properties. Second, Vietnamese lacks consonant clusters entirely; both onsets and codas are realized as single consonants, even in cases where the orthography contains multiple letters (ph, ngh, ch, ect.). Third, Vietnamese final consonants are typically unreleased (Vuong, 1998), with articulators held in place rather than released, whereas English requires full articulation of final consonants.

In the research, students themselves reported that they “find it hard” to pronounce certain clusters smoothly, often describing the experience as “getting stuck” in articulation. Such perceptions illustrate common difficulties faced by Vietnamese learners, which can be attributed to limited exposure to authentic English input, insufficient knowledge of articulatory mechanisms, and the influence of native language phonology. As a result, learners tend to rely on three main repair strategies when dealing with clusters: deletion, substitution, and vowel insertion. Additional problems occur with medial clusters, which appear between syllables or across word boundaries where learners frequently fail to produce smooth, connected speech, leading to hesitation and reduced fluency. When cluster complexity further disrupts fluency, students resort to the same three compensatory strategies.

The following sections provide a more detailed account of these challenges and outline pedagogical remedies.

A. Deletion

Sound deletion frequently occurs in final consonant clusters, functioning as a coping strategy influenced by the learners' mother tongue. In Vietnamese, codas are typically unreleased; for instance, in *cúc*, the final /k/ is held without aspiration, unlike the initial /k/. By contrast, English requires full release of both initial and final consonants, as in *cook* /kʊk/. The difficulty increases when

Problems in the Pronunciation of English Consonant Clusters – A Case Study of English-Majored Students at a Vietnamese University

clusters end in a stop consonant (e.g., /nd/, /pt/, /sp/, /skt/). Learners often omit the final stop, producing forms such as /en/ for *end*, /stop/ for *stopped*, /kris/ for *crisp*, and /as/ for *asked*.

Deletion is also observed in initial clusters that include the approximant /w/, as in *twelve* and *swim*, where /w/ is frequently dropped, resulting in /telv/* and /sim/*. In some cases, learners localize the /w/, replacing it with a glide similar to its Vietnamese counterpart, producing variants such as /toel/* or /suym/*. Likewise, the dental fricatives /θ/ and /ð/, which are absent in Vietnamese, are commonly omitted when they appear in clusters, as in *tenths* or *clothes*, which students render as /tens/* and /kləʊz/*.

Such deletions, particularly of final consonants, pose serious challenges for communication with native speakers. Because many English words are distinguished by their final consonant, accurate perception of both lexical and grammatical meaning often relies heavily on the coda, and its omission may therefore lead to misunderstanding.

B. Substitution

Vietnamese is classified as an isolating language in which syllables are produced discretely, though the internal sounds of each syllable are closely integrated. Consequently, when learners encounter consonant clusters that include sounds absent from Vietnamese, they frequently replace these forms with their native phonemes, a phenomenon that is both common and systematic. Even consonants that are produced at a similar place or manner in Vietnamese may undergo modification when combined with others, as learners adjust articulation to ease production. The following section outlines several recurring substitution patterns.

Table 4: Common substitution patterns

Sound/Cluster	Test Word	Substitution Tendency
/str/	<i>street</i>	/tr/ → /t/ → /stj:t/
/pl/	<i>plough</i>	/p/ → /f/ → /flaʊ/
/ld/	<i>world</i>	/l/ → /n/ → /wɜ:nd/
/ft/	<i>wished</i>	/f/ → /st/ → /wist/
/dʒd/	<i>judged</i>	/dʒ/ → /zd/ → /zʌdz/
/θr/	<i>three</i>	/θ/ → /t/ or /t/ → /tri:/ or /ti:/
/ŋk/	<i>think</i>	/ŋ/ → /p/ → /θɪŋk/

While these substitutions may facilitate articulation and seldom result in immediate misunderstanding, they represent non-standard forms of pronunciation. Over time, such patterns can become ingrained habits, which in turn create additional challenges in connected speech processes such as linking, elision, and liaison.

C. Vowel insertion

In contrast to deletion, some students insert a vowel (mainly a short schwa /ə/ or a short /ɪ/) into the cluster. This is a common habit among many Vietnamese who have had little pronunciation training because Vietnamese is monosyllabic and every syllable contains a vowel. In English, some consonants have vocalic properties, such as /l/. When /l/ combines with bilabials like /p/, /b/, /f/, English speakers may insert /ə/ in words such as *apple*, *table*, *beautiful* (/p(ə)l/, /b(ə)l/, /f(ə)l/). However, most clusters do not have an inserted vowel. When students pronounce clusters, they often add a very light /ə/ or /ɪ/ (marked with “+”) between consonants like /k+r/, /b+l/, /s+k/, /g+r/, so the initial consonant is lengthened: c+ream, b+lue, s+chool, g+reen. Vowel insertion is also common in final clusters, with the vowel added at the end. For example, *clothes* becomes /kləʊzə/, or *halve* becomes /hɑ:və/, dropping /l/ and adding /ə/ at the end! This not only produces non-standard pronunciation but also confuses native listeners who cannot recognize the intended syllable.

These findings are consistent with evidence from other empirical classroom and field studies. For instance, Persian-speaking EFL learners tend to insert vowels more frequently in initial clusters than in final ones, a pattern that mirrors the restrictions of Persian phonotactics (Khanbeiki & Abdolmanafi-Rokni, 2015). Similarly, Vietnamese university students often simplify clusters by deleting the first consonant in CC structures or the first or second consonant in CCC structures. They also show feature changes, especially in the voicing of aspirated plosives, while vowel insertion occurs less often (Tran & Nguyen, 2022; Hwa-Froelich et al., 2002). Taken together, such results support the Markedness Differential Hypothesis, which predicts that the more marked and less familiar a structure is in the learners’ L1, the more difficult it will be and the more likely learners are to modify it (Eckman, 1977; Anderson, 1987; Major, 1996).

D. Pedagogical Implications

For pedagogy, the literature supports several principles. First, sequence instruction by markedness: begin with two-segment stop+liquid onsets before tackling /s/+stop and tri-consonant clusters; stagger coda work from light to dense (Eckman, 1977; Vennemann, 2012). Second, treat perception and production jointly via minimal pairs and gated listening to reduce perceptual

Problems in the Pronunciation of English Consonant Clusters – A Case Study of English-Majored Students at a Vietnamese University

epenthesis (Kabak & Idsardi, 2007). Third, address timing and coordination, not just segment identity: slow-to-fast chaining, metronomic pacing, and attention to aspiration and release cues improve overlap and reduce repairs (Avery & Ehrlich, 1992; Baagbah et al., 2019). Finally, tailor feedback to L1-specific profiles (e.g., aspiration for Vietnamese; /sC/ timing in Korean; onset epenthesis for Persian/Arabic), leveraging explicit phonotactic awareness to accelerate durable gains (Choi, 2016; Tran & Nguyen, 2022).

Learners' difficulties with English consonant clusters form part of the broader challenges of pronunciation teaching and learning. Errors in cluster production largely overlap with common segmental errors, most of which can be attributed to first-language transfer and insufficient systematic practice. Addressing these issues requires coordinated efforts from both teachers and learners. At a broader, institutional level, there is a need for standardized curricula and teaching materials in pronunciation, an area that calls for sustained, long-term research and development. The present paper, however, focuses on a set of practical measures distilled from the researcher's extensive teaching experience with English-language majors at university level and the review of a wide range of pronunciation materials, both local and international.

First, students should develop a solid foundation in English phonemes through systematic practice, beginning with minimal pairs such as *bet* vs. *bed* or *plough* vs. *flour*. Useful resources include *Ship or Sheep?* (Baker, 2006) and *Pronunciation Pairs* (Baker & Goldstein, 2007). Once learners can reliably distinguish such contrasts, practice can be reinforced through online applications and multimedia tools.

Second, for consonants absent from Vietnamese, which may not necessarily cause confusion but pose articulation challenges, learners should be guided to understand their place and manner of articulation. Instruction should combine auditory models, visual diagrams of the speech organs, and guided imitation. For instance, /θ/ can be described as placing the tongue tip between the teeth and releasing airflow, while /ð/ differs only in the presence of vocal fold vibration. Descriptive explanations must be paired with repetitive, structured practice to achieve mastery.

Third, for complex clusters, the backward-building technique is particularly effective. For example, to produce /skr/ in *scream*, students can be guided progressively: /i:m/ → /ri:m/ → /kri:m/ → /skri:m/. Practice should move gradually from slow to faster tempos until the entire cluster is produced fluently. Although time-intensive, this approach trains articulators systematically and reduces production difficulty.

Different strategies can be applied depending on error type. In cases of deletion, each consonant should be practiced in isolation before moving to coarticulated sequences. For substitution, contrastive drills and minimal pairs are particularly helpful. Vowel insertion is more challenging to remediate, as it often stems from entrenched habits; sustained practice and corrective feedback are essential in addressing this issue.

Equally important is that teachers themselves develop mastery of pronunciation rules, including assimilation, elision, and connected speech processes. Consonant clusters are not only difficult for L2 learners but can also be simplified in native speech. In English phonology, processes such as assimilation and elision are governed by systematic rules. For example, plosives in clusters may be unreleased or elided, as in *asked* /ɑ:ʃt/. Similarly, assimilation occurs in sequences such as *good boy*, where /d/ assimilates to /b/, yielding /gʊb bɔɪ/. Learners must be trained to recognize when such reductions are permissible so they can distinguish rule-governed alternations from non-standard substitutions.

VI. CONCLUSIONS

Findings from this survey confirm that Vietnamese learners of English experience persistent difficulties with consonant clusters, leading to systematic use of deletion, substitution, and vowel insertion. These tendencies hinder intelligibility and pose obstacles in cross-cultural communication. Overcoming them requires collaborative effort between teachers and learners, with emphasis on mastering both phoneme inventories and the rules of pronunciation, assimilation, and connected speech. Equally critical is consistent, deliberate practice aimed at eliminating fossilized errors and replacing them with accurate habits.

Pronunciation plays a decisive role in communicative success. While learners are aware of its importance, and some engage in independent practice, errors remain frequent, particularly in cluster production. By outlining the phonological basis of these errors and suggesting pedagogical interventions, this study seeks to contribute practical insights for English-language majors at a particular university and for English learners more generally.

REFERENCES

- 1) Altakhaineh, A. R. M., Al-Junaid, H., & Younes, S. (2024). The production and spelling of English initial and final consonant clusters by Arabic-speaking EFL learners. *Languages*, 9(1), 1–17. <https://doi.org/10.3390/languages9010004>
- 2) Anderson, J. (1987). The markedness differential hypothesis and syllable structure difficulty. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 279–291). Newbury House.
- 3) Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford University Press.

Problems in the Pronunciation of English Consonant Clusters – A Case Study of English-Majored Students at a Vietnamese University

- 4) Baagbah, T. A., Jaganathan, P., & Mohamad, F. (2019). A case study to improve pronouncing English consonant clusters by young Yemeni learners in Malaysia. *International Journal of English Linguistics*, 9(1), 1–9. <https://doi.org/10.5539/ijel.v9n1p1>
- 5) Benson, B. (1986). The markedness differential hypothesis: Implications for Vietnamese speakers of English. In F. Eckman, E. Moravcsik, & J. Wirth (Eds.), *Markedness* (pp. 271–289). Springer.
- 6) Choi, J. (2016). Investigation into Korean EFL learners' acquisition of English /s/+consonant onset clusters. *Advances in Language and Literary Studies*, 7(2), 48–54. <https://doi.org/10.7575/aialc.all.v.7n.2p.48>
- 7) Cook, V. (1991). *Second language learning and language teaching*. New York: Edward Arnold.
- 8) Đoàn, T. T. (1999). *Vietnamese phonetics*. VNU Press.
- 9) Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, 27(2), 315–330. <https://doi.org/10.1111/j.1467-1770.1977.tb00124.x>
- 10) Finegan, E. (1994). *Language: Its structure and use*. Harcourt Brace College Publishers.
- 11) Gilakjani, A. P. (2011). A study on the situation of pronunciation instruction in ESL/EFL classrooms. *Journal of Studies in Education*, 1(1), 1-15.
- 12) Hwa-Froelich, D. A., Hodson, B. W., & Edwards, H. T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech-Language Pathology*, 11(3), 264–273. [https://doi.org/10.1044/1058-0360\(2002/028\)](https://doi.org/10.1044/1058-0360(2002/028))
- 13) Kabak, B., & Idsardi, W. (2007). Perceptual distortions in the adaptation of English consonant clusters. *Language and Speech*, 50(1), 23–52. <https://doi.org/10.1177/00238309070500010201>
- 14) Khanbeiki, R., & Abdolmanafi-Rokni, S. J. (2015). A study of consonant clusters in an EFL context. *International Journal of Learning, Teaching and Educational Research*, 10(4), 1–14.
- 15) Major, R. C. (1996). Markedness in second language acquisition of consonant clusters. In R. Bayley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 75–96). John Benjamins.
- 16) Ngo, P. A. (2011). L1 influence on Vietnamese accented English. *Voices*, 108, 108–125.
- 17) Nguyen, D. D. (2020). *Common errors in English speaking lessons of second-year English major students at Haiphong Technology and Management University* (Unpublished doctoral dissertation). Haiphong University of Management and Technology, Vietnam.
- 18) Roach, P. (2009). *English phonetics and phonology* (4th ed.). Cambridge University Press.
- 19) Sato, C. J. (1984). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning*, 34(4), 43–57. <https://doi.org/10.1111/j.1467-1770.1984.tb00353.x>
- 20) Tran, T. K. L., & Nguyen, A. T. (2022). Common mistakes in pronouncing English consonant clusters: A case study of Vietnamese learners. *Can Tho University Journal of Science*, 14(3), 32–39. <https://doi.org/10.22144/ctu.jen.2022.040>
- 21) Vennemann, T. (2012). Structural complexity of consonant clusters: A phonologist's view. In P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, & B. Kühnert (Eds.), *Consonant clusters and structural complexity* (pp. 11–32). De Gruyter Mouton.
- 22) Vương, H. L. (1998). *Concise Vietnamese phonetics*. Education Publishing House.
- 23) Yavas, M. (2005). *Applied English phonology*. Blackwell Publishing.



There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non Commercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.