Can Tho University Journal of Science

website: ctujs.ctu.edu.vn

Can Tho University Journal of Science

Common mistakes in pronouncing English consonant clusters: A case study of Vietnamese learners

Tran Thi Khanh Lam[1], and Nguyen Anh Thi[2*]

[1]*Student of School of Foreign Languages, Can Tho University, Viet Nam*

[2]*School of Foreign Languages, Can Tho University, Viet Nam*

*Correspondence: Nguyen Anh Thi (email: nathi@ctu.edu.vn)

Article info.

ABSTRACT

Pronunciation plays a significant role in English teaching and learning around the world and this particular field has received much attention from several scholars. The study was conducted to investigate common mistakes made by Vietnamese university learners while pronouncing English consonant clusters. This is a quantitative study with the participation of 39 English as a Foreign Language (EFL) learners. To collect data for the study, a pronunciation test was employed. The results show that the common mistakes made by the learners varied, depending on the types of consonant clusters. Specifically, clusters containing voiceless plosives led to the highest mispronunciation. There was also a tendency to simplify the complex clusters of three-consonant by deleting the first, second consonant, or in some cases, both. Accordingly, this study suggests pedagogical implications for teachers and learners in similar contexts in Vietnam in acquiring the pronunciation of English.

1. INTRODUCTION

1.1. Rationale

Pronunciation has played a significant role in English teaching and it is considered one of the most complicated skills to master among others. Therefore, learners have to invest a large amount of their time improving pronunciation (García, 2007). Gilakjani (2011) states that comprehensive pronunciation is one of the critical requirements of learners' competency and one of the most crucial factors of English learning. However, pronunciation in second language acquisition, English included, is often studied moderately due to its complexity to conduct research and its interfering variables.

Obviously, English language learners struggle a lot in studying English pronunciation (Gilakjani, 2011).

Lin (2014) claims that even if students have learned English in an English as a Second Language course, they might still lack sufficient proficiency in pronunciation. Several difficulties regarding pronunciation were found when students had to produce foreign sounds using their speech organs. Sahatsathatsana (2017) states that Thai students were more likely to pronounce English words using Thai consonant sounds. In fact, learners' first language sound patterns are likely to influence their pronunciation in target languages (Jahan, 2011; Zhang & Yin, 2009). The case is also true for Vietnamese EFL learners, especially when it comes to sound clusters found in vowels and consonants.

Some studies have revealed that pronouncing consonant clusters correctly is what many Vietnamese learners struggle with (Benson, 1986;

Sato, 1984). In this regard, Hwa-Froelich et al. (2002) conducted a study to compare and contrast between Vietnamese and English phonology to investigate whether first language (L1) transfer occurred when pronouncing English words. The results showed that Vietnamese EFL learners tended to have problems with English suprasegmental and segmental features, especially with final consonant sounds (Ngo, 2011; Nguyen, 2020). However, in Vietnam, it is said that research on consonant clusters is scarcely conducted. Thus, to fill the gap in the literature in the field of pronunciation, this study was conducted to investigate common mistakes that Vietnamese EFL learners often make when pronouncing certain English consonant clusters. To this end, the present study addresses the following research question: What are the common mistakes Vietnamese EFL learners make when pronouncing English consonant clusters?

1.2. Literature review

1.2.1. Issues in English pronunciation: An overview

Comprehensible pronunciation is one of the requirements of language learners' competence and it is also one of the most significant factors of language instruction (Gilakjani & Sabouri, 2016). While good pronunciation is said to lead to mastering a language, poor pronunciation could lead to considerable difficulties in language learning (Gilakjani, 2011).

Foreign contexts. The issue of mispronunciation in language learning has occurred as a problematic phenomenon for all language learners, particularly English ones when they try to master English as a foreign language. Cruz (2005) showed how minimal pairs could illustrate meaning confusion derived from mispronunciation in Brazilian learners when producing English. Tsuzuki and Nakamura (2009) conducted a study in which they investigated the English intelligibility of Japanese learners. The results showed that three types of mispronunciation were found to seriously hinder intelligibility including mispronunciation in consonants (such as plosives and liquids), alternation of vowel lengths, and misplacement or absence of word stress. Similarly, Meng et al. (2010) compared phonological systems between the two languages of Cantonese and English. The results revealed certain differences in phonology between the two researched languages. For example, a native Cantonese speaker's second language (L2) speech frequently replaces the voiced fricative /v/ with an unvoiced fricative /f/. Therefore, mispronunciation in any form due to linguistic differences among languages could lead to significant misunderstandings (Kashiwagi & Snyder, 2008).

The Vietnamese context. Mispronunciation in English has occurred as a big concern among Vietnamese learners of English and this has therefore received great attention from a large number of researchers. Ha (2005) conducted a study examining how Vietnamese learners pronounced English consonant sounds while communicating in English. The author concluded that Vietnamese students tended to skip most consonant sounds that occurred in the ending position since it seemed to be difficult for them to generate sounds like /ʒ/, /ʤ/, and /ʧ/ in the ending positions. Nguyen (2020) conducted a study in which the author employed a contrastive analysis method to investigate Vietnamese learners' difficulty in pronouncing the final English consonant sounds. The results indicated that most Vietnamese learners encountered problems while pronouncing these sounds since there were certain differences between Vietnamese and English, that was to say, while English has final consonant sounds, Vietnamese does not.

A comparison between the two languages of English and Vietnamese concerning the phonological systems has also been made. Ngo (2011) conducted a study comparing vowels and consonants between Vietnamese and English which concluded that the consonants of the two languages are considerably different. Specifically, the /z/ sound in the Vietnamese language is different from its English counterpart. In addition, common mistakes made by Vietnamese learners when trying to speak English have been researched. Nguyen (2008) highlighted that Vietnamese learners of English tended to make mistakes in pronunciation including leaving out ending sounds that seemed too strange or challenging for them, reducing final consonants and clusters and moving back to their original language, and adding schwa sound in the last clusters. Also, Duong (2009) found 4 pairs of confusing consonants that were often mistaken by Vietnamese learners when pronouncing English words. They were /ʃ/, /ʒ/, /ʧ/ and /ʤ/. The author also addressed four possible reasons explaining for making these mistakes including (1) failure in distinguishing the differences; (2) influence of the mother tongue; (3) perceptions of mistakes and (4) inadequate drills. In addition, Nguyen (2019) summarized two main types of common English errors of Vietnamese

learners while pronouncing English including the frequent omission of ending sounds and sound confusion (some English sounds do not have their Vietnamese counterparts).

1.2.2. *Consonant clusters in the English language*

A consonant cluster is a group of consonants that have no intervening vowel (Vennemann, 2012). In the English language, consonant clusters can be in the positions of onset or coda within a syllable. Initial consonant clusters, known as onsets, comprise two or three consonants that occur at the beginning of a syllable. They are referred to as two-consonant clusters, such as Storm /stɔːrm/, Umbrella /ʌmˈbrelə/, and three-consonant clusters, such as *String* /strɪŋ/, *Display* /dɪˈspleɪ/. Final consonant clusters, known as codas, contain up to four consonants and occur at the end of a syllable, for instance, *end* /end/ or *text* /tekst/. In what follows, only consonant clusters in the onset or initial position are discussed since this type is the main focus of the present study.

Initial consonant clusters are of two types: initial two-consonant clusters (CC) and initial three-consonant clusters (CCC). According to Roach (2009), the initial two-consonant clusters consist of two kinds. One is composed of a pre-initial /s/ followed by one initial consonant. The other begins with one initial consonant of a set of fifteen consonants, followed by one post-initial consonant of the set /l/, /r/, /w/, and /j/. Initial three-consonant clusters always begins with a pre-initial /s/, followed by one initial consonant of the set /p/, /t/, /k/, and one post-initial consonant of the set /l/, /r/, /w/, and /j/.

These complex structures of onsets and codas often cause difficulty for learners whose native languages do not allow initial or final consonant clusters (Avery & Ehrlich, 1992). The languages investigated in many studies have been ones that either allow no consonant clusters in any position (Vietnamese and Mandarin) or allow certain consonant clusters in word-final position only (Arabic), or word-initial position only (Portuguese). In this respect, learners from different language backgrounds seem to have different preferences for cluster simplification, ranging from vowel epenthesis (adding an extra syllable), and substitution (replacing a sound), to deletion or feature changes in certain segments in the clusters.

While the L1 Mandarin and Arabic speakers seem to prefer deletion (Anderson, 1987), the Portuguese natives' most common strategy is a substitution of the voicing contrast (Major, 1996). In the case of Vietnamese EFL learners, Sato (1984) did not find epenthesis to be a common strategy used in the pronunciation of consonant clusters while Benson (1986) found a considerable preference for insertion of schwa after obstruents, suggesting an optimal regression to the basic Consonant Vowel structure.

1.2.3. *Factors affecting learners' pronunciation*

Different factors affecting learners' pronunciation have been identified in different studies. L1 is one of the main causes of learners' mispronunciation. In fact, the differences among languages would make learners struggle to pronounce to a certain extent. The Markedness Differential Hypothesis formulated by Eckman (1977) describes how the differences in features between L1 and L2 might be difficult for learners to acquire. Pertaining to consonant clusters which cannot be found in many Asian languages such as Vietnamese, Japanese, and Korean. Theseare a real challenge for learners who are native speakers of these languages and they themselves struggle to master. Choi (2016) states that syllable structures of languages are one of the significant examples of cross-linguistic differences. This idea is based on the concept that the transfer occurring between L1 and L2 has been one of the most important issues for linguists, and has been regarded as the major source of difficulties for L2 learners (Yavas, 2005). Choi (2016) argues that two or more consonants occurring at the onset of a syllable simply do not exist in Korean, but English allows two or three consonants at the beginning of syllables, which has created a true obstacle for Korean learners when learning English. Ioup and Weinberger (1987) suggested that one of the main sources of difficulties in interlanguage phonology is the negative transfer from the native language (VanPatten & Houston, 1998). Avery and Ehrlich (1992) also consider L1 as a crucial factor, showing that the sound pattern of the learners' L1 transferred into the L2 and is likely the cause of foreign accents, which then reduce L2 speaker intelligibility.

Aptitude for oral mimicry or a considerable amount of time spent in an English-speaking country could influence the speech production of L2 speakers (Purcell & Suter, 1980). Skehan (2002) argues that aptitude batteries should capture the abilities drawn on at different stages of L2 processing, including phonetic sensitivity, grammatical sensitivity, paired associates, and rote memory abilities. Carroll (1981) states that four factors are related to the ability to

learn languages such as phonemic coding ability (memory of sounds and their combinations), associative memory (the ability to remember new words), inductive language learning ability (the ability to find patterns in words and sentences), and grammatical sensitivity (the ability to understand sentence structure of unknown languages). Needless to say, learners' aptitude plays an important role in helping them master L2 pronunciation, especially sounds that are not included in their native language system.

2. MATERIALS AND METHOD

2.1. Design and participants

A quantitative method was employed in this present study. 36 Vietnamese EFL students at a university in Vietnam volunteered to participate in this study. There were 5 males (13.9%) and 31 females (86.1%) with the age range from 19 to 20 years. All of the participants have been studying English for at least 7 years at secondary school and 1 year at university. All the participants were reported to have no speech difficulties, hearing difficulties, or any other disabilities that may affect the results of the study. In addition, they have never been to any of the countries where English is spoken as the first language. Therefore, they do not have any kind of exposure to a native English environment. Prior to data collection, informed consent was obtained from all participants.

2.2. Instrument

A pronunciation test comprising three sub-tests of (1) a wordlist test, (2) a question-response test, and (3) a long-speech test was employed in this study. To be specific, the pronunciation test consists of ten English consonant clusters including five two-consonant (CC) clusters of /br/, /kr/, /pr/, /st/, /tr/ and five three-consonant (CCC) clusters of /skr/, /skw/, /spl/, /spr/, /str/. It is noted that these ten English consonant clusters were selected based on the list of most occurring English consonant clusters in the onset positions, which was proposed by Duanmu (2008). In what follows, a detailed description of the three sub-tests will be presented.

Wordlist test

The wordlist test consisted of 30 words that each included one of the ten consonant clusters mentioned above. The words were presented in 10 sets of three, where the words in the same set had the same consonant cluster in the onset position. The participants were asked to pronounce each word

twice allowing their pronunciation of a particular English consonant cluster to be measured.

Question-Response test

The question-response test consisted of 10 questions with 10 illustrated pictures for response suggestions. Each picture was meant to illustrate one word which contained one consonant cluster from the ten consonant clusters targeted in the test. The participants were then asked to take turns to give the answers using the words suggested in the illustrated pictures.

Long-speech test

In this test, 10 long sentences containing the ten targeted consonant clusters of the study were employed. Each sentence was designed to consist of 2 words representative of one type of consonant cluster. Doing so allows the researcher to accurately record participants' pronunciation performance. Each sentence was approximately 13 words long, but data were collected only from the targeted words within the sentence (i.e., words that contained the consonant clusters tested). The participants were asked to read aloud each sentence twice. Table 1 below summarizes data on the types of consonant clusters that appeared in the tests.

Table 1. Types of consonant clusters appeared in the tests

Type of Consonant Clusters	Wordlist	Question-Response	Long-Speech
/br/	3	1	2
/kr/	3	1	2
/pr/	3	1	2
/st/	3	1	2
/tr/	3	1	2
/skr/	3	1	2
/skw/	3	1	2
/spl/	3	1	2
/spr/	3	1	2
/str/	3	1	2

Identifying pronunciation mistakes

There were two main types of consonant clusters (CC and CCC) in the onset positions being tested. Each type consisted of 5 clusters (Table 1) and was tested 6 times throughout the pronunciation test. Totally, 1080 tokens were counted for identifying students' mistakes in pronouncing English consonant clusters in each type of consonant cluster.

To collect data on students' mistakes, 36 recordings from the participants, as well as scoring rubrics were

sent to 2 independent examiners for scoring purposes. Each examiner listened to each recording, and carefully noted down the words containing the consonant clusters that were incorrectly pronounced. Then, the incorrect items were classified into 3 main categories of Deletion (D), Epenthesis (E), and Feature Change (FC). To ensure reliability, Pearson's correlation coefficient was computed to assess the interrater reliability between the two examiners (see Table 2). The results showed that a strong correlation was found (r_{test} =0.98; n=30; p<.001).

Table 2. Pearson's correlation between the two independent examiners

	Correlations	
		Pronunciation Test (2nd examiner)
Pronunciation Test (1st examiner)	Pearson Correlations	0.98**
	Sig. (2-tailed)	.000
	N	30

*** Correlation is significant at the 0.01 level (2-tailed).*

2.3. Procedures

Prior to the study, the researcher firstly contacted the potential participants and invited them to volunteer to participate in the study. Then, the participants were informed of the objectives and aims of the research, and their identities were kept confidential. Soon after all participants had been eligibly selected, they were asked to complete a consent form and then arranged an online meeting for data collection. All of the participants were introduced to the test format and instructed on how to do the test beforehand. In addition, to ensure the reliability, validity, and practicality of the test, the researcher conducted a pilot test on a group of 10 students of similar characteristics. The piloting activity allows the researcher to determine the feasibility of the test, as well as to improve the reliability of the test items.

The tests were conducted online via Zoom and video and audio-recorded for data analysis afterwards. Each participant spent approximately 10 minutes completing the test. Before officially taking the test, the examiner explained and gave detailed instructions on how to deal with the test. To be specific, regarding part 1 of the pronunciation test, the Wordlist test, the participants were asked to pronounce the word appearing on the screen twice. The purpose of this test was to test the participants'

ability to pronounce a single tested word in which a particular English consonant sound was intentionally included. In the second part, the Question-response test, the participants were asked to answer the questions appearing on the screen with the illustrated pictures. The participants were required to pronounce the words which contained the consonant clusters when putting them in a question-answer context. In the last part, the Long-speech test, the participants were asked to read each sentence appearing on the screen twice with their normal reading speed. The purpose of this test was to evaluate the participants' ability to pronounce words precisely when putting them in a long sentence in which suprasegmental features could influence the production of consonant clusters in the words.

2.4. Data analysis

The data of the study were quantitatively analyzed. All test data were analyzed for descriptive statistics to identify learners' common mistakes in pronouncing English consonants with regard to CC and CCC clusters.

3. RESULTS AND DISCUSSION

Table 3 summarizes the descriptive statistics of participants' overall results in pronouncing the two types of consonant clusters of CC and CCC, categorized by Expected Productions (correct pronunciation) and Mistakes (incorrect pronunciation). The results generally showed that the participants performed quite well on the pronunciation test for both types of consonant clusters with the percentage of Expected Productions being 76.67% and 83.89% respectively compared to Mistakes (23.33%, 16.11%, respectively).

Table 3. Descriptive statistics of participants' pronunciation test

Clusters	Productions	Tokens (n=1080)	Percentage (%)
CC	Expected Productions	828	76.67%
	Mistakes	252	23.33%
CCC	Expected Productions	906	83.89%
	Mistakes	174	16.11%

3.1. Learners' mistakes in pronouncing CC clusters

Regarding learners' mistakes in pronouncing CC clusters which were categorized into D, E, and FC,

the results (see Table 4) indicated that FC was found to be the most frequent type of mistake made by the students accounting for 77.4% of the total, followed by D and E (22.2%, 0.4%, respectively). In addition, with regard to D, the results showed that the participants tended to delete the first consonant in the CC clusters (55 out of 56 recorded tokens) instead of the second one (1 out of 56 recorded tokens). The results also indicated that there was little E found in the production of CC clusters (0.4%).

Table 4. Learners' mistakes in the production of CC clusters

Productions	Recorded Tokens	Percentage (%)
Mistakes	252	100%
Deletion	*56*	*22.2%*
– *1ˢᵗ consonant*	*55*	
– *2ⁿᵈ consonant*	*1*	
Epenthesis	*1*	*0.4%*
Feature Change	*195*	*77.4%*

3.2. Learners' mistakes in pronouncing CCC clusters

Regarding learners' mistakes in pronouncing CCC clusters, the results (see Table 5) showed that D took the lead, with 78.2% of the total mistakes belonging to this category. Specifically, the results revealed that the learners tended to delete the first consonant (77 out of 136 recorded tokens) and the second consonant (56 out of 136 recorded tokens) in the CCC clusters. However, dealing with learners' deletion of the third, and the first&second consonant in the CCC clusters, it is noted that only a small number of mistakes was found in cases of the third (1 out of 136 recorded tokens) and the first&second (2 out of 136 recorded tokens). In addition, FC was found to be the second most common mistake type in the learners' production of CCC clusters with 16.6%, followed by E which accounted for 5.2%.

Table 5. Learners' mistakes in the production of CCC clusters

Productions	Recorded Tokens	Percentage (%)
Mistakes	174	100%
Deletion	*136*	*78.2%*
– *1ˢᵗ consonant*	*77*	
– *2ⁿᵈ consonant*	*56*	
– *3ʳᵈ consonant*	*1*	
– *1ˢᵗ and 2ⁿᵈ consonant*	*2*	
Epenthesis	*9*	*5.2%*
Feature Change	*29*	*16.6%*

3.2. Discussion

Overall, the results (see Table 4) collected from learners' performance on CC clusters showed that FC was found to be the most frequent type of mistake made by the learners. A possible reason for this could be explained that the learners found it difficult in pronouncing CC clusters; and therefore, they tended to change voiceless plosives with an approximant in CC clusters, such as /t/, /p/, and /k/ in /tr/, /pl/, and /kr/. Another reason for FC to be the most frequent type of mistake could be seen in the interference of learners' L1 in the production of voiceless sounds (Nguyen, 2021). In fact, according to Language and Orientation Resource Center (1981), aspiration in consonants does not commonly occur in Vietnamese phonology. Vietnamese allows slight aspiration in an alveolar such as tʰ, but not in other positions such as bilabial or glottal. Also, a complex combination of a voiceless plosive and an approximant is not commonly seen in the Vietnamese language (Nguyen & Dutta, 2017). Therefore, the learners had a tendency to change the voicing feature of voiceless consonants into the voiced ones (Hwa-Froelich et al., 2002). It is implied that Vietnamese learners should pay attention to aspirated sounds such as /p/, /k/, or /t/ when pronouncing English words since the voicing features between the two languages are different.

With regard to D, the results showed that the learners tended to delete the first consonant in the CC clusters instead of the second one. Among the five CC clusters tested, although not being intentionally selected, the second consonant of four of them was the approximant /r/. Since the Vietnamese approximant /r/ is similar to the English /r/ in terms of pronunciation, the learners had no difficulty pronouncing it. However, the combination with a voiceless plosive preceding tends to be difficult for Vietnamese students (Nguyen, 2021). Those who did not change the voicing feature of the voiceless plosives would opt for deletion in these sounds, creating cluster reduction in the first consonant.

The results also indicated that there was little E found in the production of CC clusters. We found this result interesting since it is slightly different from what Sato's study (1984) reported, which indicated that there was no E found in the Vietnamese learners' pronunciation of consonant clusters. The result of the current study revealed that the insertion of schwa after obstruents such as /b/ would help simplify the cluster and ease its articulation.

Regarding learners' mistakes in pronouncing CCC clusters, the results (see Table 5) indicated that D was found to be the most common type of mistake, followed by FC and E. This can be explained by the Markedness Differential Hypothesis proposed by Eckman (1977) which suggests that L2 learners may have more difficulty acquiring the features of the target language that are different from their mother tongue. In fact, Vietnamese phonology does not allow clusters that contain more than 3 different consonants. Therefore, clusters of three consonants seem to be a challenge for the L2 learners in the study (Bouchhioua, 2019). As a result, the learners resorted to D in order to break down and simplify the clusters. Furthermore, the results showed that the learners tended to delete the first and the second consonant (with a total of 133 out of 136 tokens) instead of the third one. This is because the pre-initial /s/ does not appear in any consonant clusters in the Vietnamese language. Therefore, adding it to the already complex two-consonant clusters is a challenge for Vietnamese learners. In a similar vein, the results also indicated that learners, to some extent, tended to insert an epenthetic vowel in order to simplify the clusters. It is therefore suggested that Vietnamese EFL teachers should focus on practicing clusters for learners. Helping them recognize the syllables within a word might be useful in terms of E. Although cluster simplification might not affect the comprehensibility of L2 speakers (Bouchhioua, 2019), it partly contributed to their classification as having a strong accent.

In terms of FC, the results showed that learners also attempted to change the correct pronunciation of the consonant sounds due to difficulty that they might encounter while aspirating plosives such as /p/ or /k/ in the middle of the three-consonant clusters (Hwa-Froelich et al., 2002). Hence, they opted for feature changing, specifically in voicing, in order to ease the articulation. Thus, it is implied that Vietnamese EFL teachers should notice that plosives similar to both languages may differ based on aspiration styles. Therefore, it is inadvisable to teach L2 learners of English the pronunciation of certain sounds interchangeably with their mother tongue.

4. CONCLUSIONS

This study investigated common mistakes in the production of English consonant clusters made by Vietnamese EFL learners. The results suggest that the common mistakes varied, depending on the types of consonant clusters. These findings were consistent with those reported from previous studies (e.g., Sato, 1984; Nguyen, 2008), in terms of consonant clusters production. However, this study has specifically focused on the mistakes in consonant clusters in the onset position.

As far as the L1 transfer is concerned, given the fact that consonant clusters are not common in the Vietnamese language, evidence from this study suggests that the participants may have transferred some aspects of their L1 phonology into their acquisition of English. Specifically, clusters containing voiceless plosives led to the highest mispronunciation.

This study has provides insightful views related to EFL phonology. The analysis presented here may prove to be beneficial not only for Vietnamese EFL teachers and learners but also for EFL teachers and learners from other similar language backgrounds.

Nevertheless, there are some limitations of this study that need to be pointed out. First, the number of participants in the study was quite small, which significantly limits the generalizability of the findings, especially in terms of the common mistakes. Second, this study focused on investigating mistakes in pronouncing English consonant clusters of learners at university. Therefore, more groups of learners such as high school learners or children should be investigated so as to set a clearer picture of Vietnamese learners with regard to English pronunciation-related issues.

REFERENCES

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.

García, A. C. (2007). The role of phonetic training in L2 speech learning. In *Phonetics Teaching & Learning Conference (PTLC)*.

Avery, P., & Ehrlich, S. (1992). Teaching American English: Oxford handbook for language teachers.

Benson, B. (1986). The markedness differential hypothesis: Implications for Vietnamese speakers of English. In *Markedness* (pp. 271-289). Springer.

Bouchhioua, N. (2019). Epenthesis in the production of English consonant clusters by Tunisian EFL learners, *Applied Linguistics Research Journal*, *3*(4), 33–44.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude*, (p.83-118). Newbury House.

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.

Cruz, N. (2005). Minimal pairs: Are they suitable to illustrate meaning confusion derived from mispronunciation in Brazilian learners' English. *Linguagem & Ensino*, *8*(2), 171-180.

Duanmu, S. (2008). Syllable structure: How different can it be in human languages.

Duong, T. N. (2009). Mistake or Vietnamese English. *VNU Journal of Foreign Studies*, *25*(1), 41-50.

Eckman, F. R. (1977). Markedness and the Contrastive Analysis Hypothesis. *Language Learning*, 27(2), 315–330.

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.

Gilakjani, A. P., & Sabouri, N. B. (2016). Why is English pronunciation ignored by EFL teachers in their classes? *International Journal of English Linguistics, 6*(6), 195.

Ha, C.T. (2005). Common pronunciation problems of Vietnamese learners of English. *VNU Journal of Foreign Studies*, *21*(1), 35-46.

Hwa-Froelich, D., Hodson, B. W., & Edwards, H. T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech-Language Pathology*, *11*(3), 264-273.

Ioup, G., & Weinberger, S. (1987). *Interlanguage phonology: The acquisition of a second language sound system*, *8*. Newbury House.

Jahan, N. (2011). Teaching and learning pronunciation in ESL/EFL classes of Bangladesh. *Journal of Education and Practice*, *2*(3), 36-45.

Kashiwagi, A., & Snyder, M. (2008). American and Japanese listener assessment of Japanese EFL speech: Pronunciation features affecting intelligibility. *Journal of Asia TEFL*, *5*(4), 27-47.

Language and Orientation Resource Center. (1981). *Refugee Education Guide general information series #23, Teaching English to speakers of Vietnamese*. Center for Applied Linguistics.

Lin, C. Y., Wang, M. I. N., Idsardi, W. J., & Xu, Y. I. (2014). Stress processing in Mandarin and Korean second language learners of English. *Bilingualism: Language and Cognition*, *17*(2), 316-346.

Major, R. C. (1996). Markedness in second language acquisition of consonant clusters. *Second Language Acquisition and Linguistic Variation*, 75-96.

Meng, H., Lo, W. K., Harrison, A. M., Lee, P., Wong, K. H., Leung, W. K., & Meng, F. (2010). Development of automatic speech recognition and synthesis technologies to support Chinese learners of English: The CUHK experience. *Proc. APSIPA ASC*, 811-820.

Ngo, P. A. (2011). L1 influence on Vietnamese accented English. *Voices*, 108-125.

Nguyen, D. D. (2020). *Common errors in English speaking lessons of second-year English major students at Haiphong Technology and Management University* (Doctoral dissertation, Đại học Quản lý và Công nghệ Hải Phòng).

Nguyen, H. D. (2019). Errors in word-final consonant pronunciation in Vietnamese English interlanguage. In *Proceedings of 15th International Conference on Humanities and Social Sciences*.

Nguyen, H. T., & Dutta, H. (2017). The adaptation of French consonant clusters in Vietnamese phonology: An OT account. *Journal of Universal Language*, *18*(1), 69-103.

Nguyen, T. H. (2021). Contrastive Analysis of Consonants in English and Vietnamese. *Journal of English Language Teaching and Applied Linguistics*, *3*(6), 58-65.

Nguyen, N. (2008). Interlanguage phonology and the pronunciation of English final consonant clusters by native speakers of Vietnamese. *Unpublished Master's thesis. Ohio University*.

Purcell, E. T., & Suter, R. W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, *30*(2), 271-287.

Roach, P. (2009). *English Phonetics and Phonology: A Practical Course*. Cambridge University Press.

Sahatsathatsana, S. (2017). Pronunciation problems of Thai students learning English phonetics: a case study at Kalasin University. *Journal of Education, Mahasarakham University*, *11*(4), 67-84.

Sato, C. J. (1984). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning*, *34*(4), 43-58.

Skehan, P. (2002). Theorizing and updating aptitude. *Individual Differences and Instructed Language Learning*, *2*, 69-94.

Tsuzuki, M., & Nakamura, S. (2009). Intelligibility assessment of Japanese accents. *T. Homann*.

VanPatten, B., & Houston, T. (1998). Contextual effects in processing L2 input sentences. *Spanish Applied Linguistics*, *2*(1), 53-70.

Vennemann, T. (2012). Structural complexity of consonant clusters: A phonologist's view. In *Consonant clusters and structural complexity*, 11-32. De Gruyter Mouton.

Yavas, M. (2005). *Applied English phonology*. Blackwell Publishing.

Zhang, F., & Yin, P. (2009). A study of pronunciation problems of English learners in China. *Asian Social Science, 5*(6), 141-14.