

Vietnamese EFL learners' Difficulties with English Consonants

Luu Trong Tuan^{1,*}

¹Ho Chi Minh City University for Natural Resources and Environment and Hung Vuong University, Vietnam

*Corresponding author.

Address: Ho Chi Minh City University for Natural Resources and Environment and Hung Vuong University, Vietnam

Received 11 June 2011; accepted 27 August 2011

Abstract

The research was conducted to examine the most problematic English consonants facing the students at Hung Vuong University. Data collated from the questionnaire survey reveals that the students have the most difficulty in pronouncing the English fricatives /ʃ/, /ʒ/ and the affricates /tʃ/ and /dʒ/ among the English consonants surveyed. The students tend to substitute Vietnamese sounds for the English sounds.

Key words: English consonants; Fricatives; Affricates; EFL

Luu Trong Tuan (2011). Vietnamese EFL learners' difficulties with English Consonants. *Studies in Literature and Language*, 3(2), 56-67. Available from: URL: <http://www.cscanada.net/index.php/sll/article/view/j.sll.1923156320110302.434>
DOI: <http://dx.doi.org/10.3968/j.sll.1923156320110302.434>

INTRODUCTION

Having been a teacher of English for several years and from my own experience of teaching English pronunciation, phonology, listening and speaking skills, public speaking subject and so on, the author of this thesis has found that students usually have difficulty pronouncing a certain number of the English consonants since the students' English accent has been strongly affected by Vietnamese. When they encounter the English consonants, they tend to substitute the Vietnamese consonants for the English ones. This process is a natural

process in learning a foreign language of adult learners. Thus, teachers of English should teach the students the manner of articulation, the place of articulation, kinds of consonants, voicing of both English consonants (the target sounds) and the Vietnamese sounds (the sounds that Vietnamese students tend to use to substitute for the English consonants). On the surface, some of the characteristics of English consonants are similar to the Vietnamese ones. However, they internally differ from each other a lot.

The four English consonants such as /ʃ/, /ʒ/, /tʃ/ and /dʒ/ do not really exist in Vietnamese. On the surface, some of the characteristics of the English consonants /ʃ/, /ʒ/, /tʃ/ and /dʒ/ are similar to the Vietnamese consonants. However, they internally differ from each other a lot. For example, in the case of the English consonant /tʃ/ and the Vietnamese consonant /c/: First, the Vietnamese consonant /c/ as in "cho" (meaning "give") can occur in word-initial position and in word-final position while the English consonant /tʃ/ can occur in word-initial position, word-medial position and word-final position. Second, the Vietnamese consonant /c/ is voiceless, laminal, palatal stop and the English /tʃ/ is voiceless, laminal, post-alveolar affricate (Roach, 1990, p.62). Third, the Vietnamese /c/ is stop, whereas the English affricate /tʃ/ is "a combination of the articulatory characteristics of a stop and a fricative". Moreover, the Vietnamese stop /c/ in word-final position is "pronounced but unreleased" but the English affricate /tʃ/ is pronounced and released in all positions.

This research was guided by the following question: What problems do the EFL students at Hung Vuong University face in acquiring /ʃ/, /ʒ/, /tʃ/ and /dʒ/?

1. LITERATURE REVIEW

1.1 Speech Sounds

“Vocal sounds are produced in the human body by the organs of what is called the vocal tract” which consists of “the nasal cavity and the oral cavity”. (Richards, Platt and Platt, 1992, p.401) and “vocal sounds which are organized to communicate information are called speech sounds.” In other words, speech sounds are made by air moving outward from the lungs through the mouth or nose” (Avery and Ehrlich, 1995, p.10). The main purpose of the speech sounds is produced by the vocal organs for the sake of communication.

According to Ladefoged and Maddieson (1996, p.12), the world’s languages contain two classes of speech sounds such as consonants and vowels and each language in the world has the different number of consonants and vowels.

As for the production of a vowel, “a speech sound in which the airstream from the lungs is not blocked in any way in the mouth or throat, and which is usually pronounced with the vibration of vocal cords.” All vowels and a number of the English consonants are voiced.

In order to produce an English consonant, an English speaker has to use the airstream from the lungs “is completely blocked (stops: /p/, /b/, /t/, /d/), partially blocked (lateral: /l/) or where the opening is so narrow that the air escapes with audible friction (fricative: /f/, /v/, /θ/, /ð/, /s/, /z/, /ʃ/, /ʒ/). With some consonants (NASALS) the airstream is blocked in the mouth but allowed to escape through the nose”(Richards, Platt and Platt, 1992, p.78). Therefore, their voicing depends on the way that each sound is produced.

1.2 Articulatory Features

The English consonants can be judged from the three main characteristics: voicing, manner of articulation and place of articulation (Roach, 1990, p.62) and (Fromkin, Rodman, Collins & Blair, 1988, p.58). These characteristics help the world’s linguists describe the individual consonants and distinguish them from each other.

1.2.1 Voicing

The first way in which consonants can be different from each other is in terms of voicing. The English consonants can be “voiced or voiceless.” (Rogers, 2000, p.47) The voicing of consonants depends on the state of the glottis, which is “the space between the vocal cords.” (Richards,

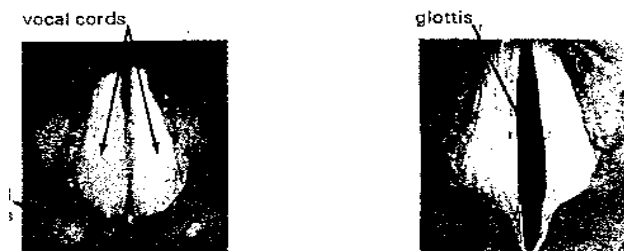


Figure 1
Two States of the Glottis (cited from Ladefoged, 2000, p.123)

Platt and Platt, 1992, p.401) (see Figure 1).

In a voiced sound, “the vocal folds are close together and vibrating.” In other words, “speech sounds which are produced with the vocal cords vibrating are called voiced. Such vibration can be felt when touching the neck in the region of the larynx” (Richards, Platt and Platt, 1992, p.402). In a voiceless sound, “the vocal cords are apart and without vibration”. Therefore, “there is little or no airflow through the glottis” as in Figure 1 (Ladefoged, 2000, p.123).

Based on the above criteria and according to Roach (1990, p.62), the sounds /f/ and /tʃ/ are voiceless and the sounds /z/ and /dʒ/ are voiced in English.

1.2.2 Manner of Articulation

The manner of articulation is the way in which speech sounds are produced by the speech organs such as oral stops, fricatives, affricates, nasal stops, a lateral and approximants (Fromkin, Rodman, Collins & Blair, 1988, p.58). However, this thesis only focuses on the two English fricatives /f/, /z/ and the two affricates /tʃ/, /dʒ/ only.

Firstly, a fricative is “a speech sound produced by allowing the airstream from the lungs to escape with friction, which is caused by bringing the two articulators, e.g. the upper teeth and the lower lip close together but not close enough to stop the air stream completely” (Richards, Platt and Platt, p.148). For example, /f/ as in “sheet”, “washing”, “fish”, and /z/ as in “genre”, “pleasure” and “garage”.

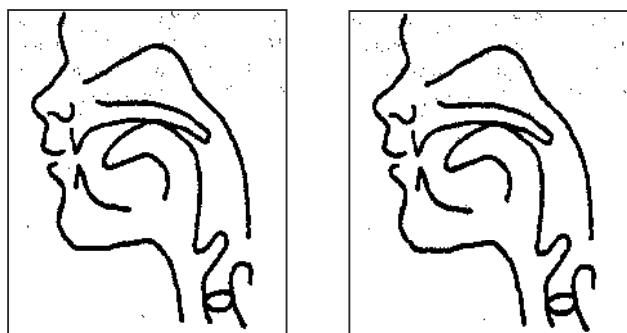


The airflow coming out from the lungs is shown by the arrows.

Figure 2
The Illustration of the English Sound /f/ and /z/ (Yates, 2003, p.92, 94)

Secondly, an affricate is “a speech sound produced by stopping the airstream from the lungs, and then slowly releasing it with friction. The first part of an affricate is similar to a stop and the second part is similar to a fricative” (Richards, Platt and Platt, 1992, p.11). The oral stops in the English affricates /tʃ/, /dʒ/ are the English /t/ and /d/. To produce an oral stop, speakers have to stop the airflow from the lungs completely and then releasing it suddenly. For example, the voiceless affricate /tʃ/ as in “cheat” /tʃi:t/ and the voiced affricate /dʒ/ as in “just”

/ˈdʒʌst/. As for the second part, they are fricatives which are above-mentioned.



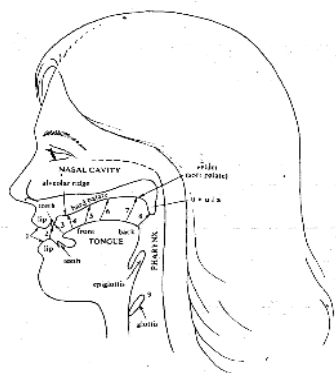
The closing phrase of /t/ in /tʃ/ and /d/ in /dʒ/ are shown by the arrows

Figure 3
The Illustration of the English Sounds /tʃ/ and /dʒ/
(Yates, 2003, p.88, 90)

In short, the manner of the English fricatives /ʃ/, /ʒ/ and the English affricates /tʃ/, /dʒ/ is the term referred to the way the airstream from the lungs coming out through the oral cavity. The airflow of the English fricatives /ʃ/ and /ʒ/ moves through an obstacle course which is created by different configuration of different speech organs. As for the oral stops /t/ and /d/ in the English affricates /tʃ/ and /dʒ/, the airflow coming out from the lungs is completely stopped and then /t/ and /d/ move to the next sounds /ʃ/ and /ʒ/. There is “no release” between /tʃ/ and /dʒ/. These characteristics are just a part of producing each of the said English sounds.

1.2.3 Place of Articulation

In production of sounds, “air passes through one or both of two passageways: the oral cavity (mouth) or the nasal passageway (nose), depending on whether the nasal passageway is blocked off or not” (Celce-Murcia, Brinton and Goodwin, 2002, p.42). The point where the contact of the articulators is called place of articulation (see Figure 4).

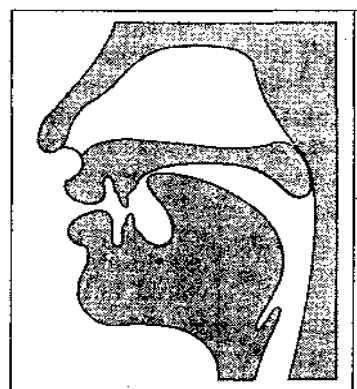


1. bilabial; 2. labio-dental; 3. dental or interdental; 4. alveolar; 5. palato-alveolar; 6. palatal; 7. velar; 8. uvular; 9. glottal

Figure 4
Places of Articulation (cited from Fromkin, Rodman, Collins and Blair, 1990, p.31)

Here are the descriptions of the place of articulation of the English fricatives /ʃ/, /ʒ/ and the English affricates /tʃ/, /dʒ/.

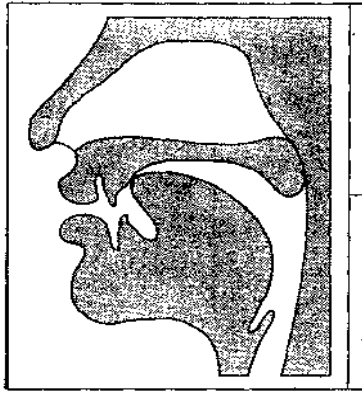
Firstly, the place of the articulation of English fricatives /ʃ/ and /ʒ/ is post-alveolar. A post-alveolar sound is a speech sound articulated with “the tongue blade makes light contact with the alveolar ridge and the soft plate is raised” so that the airflow only escapes through the oral cavity (Kelly, 2000, p.51). In other words, a post-alveolar sound is produced with the tongue near or touch the back of the alveolar ridge, further back in the mouth than alveolar consonants. The difference between the English /ʃ/ and the English /ʒ/ is the voicing only.



The place of articulation of /ʃ/ and /ʒ/ is marked by an arrow (Kelly 2000, p.51)

Figure 5
The Illustration of the Place of Articulation of /ʃ/ and /ʒ/

Secondly, the place of articulation of the English /tʃ/ and /dʒ/ is also post-alveolar. But they are much more complex than the English /ʃ/ and /ʒ/ because the English affricates consist of oral stops and fricatives. As for the plosives /t/ in /tʃ/ and /d/ in /dʒ/, they are produced with “tongue blade against the alveolar ridge and the soft plate is raised” so that the airflow can escape through the oral cavity instead of the nasal cavity (Kelly, 2000, p.49). Not only that, according to Roach (1990, p.52), the /t/ component in the English affricates /tʃ/ and /d/ in /dʒ/ have the place of articulation rather further back in the mouth than /t/ and /d/ usually has. For example, /ʃ/ in /tʃ/ as in “chance” /ˈtʃɑːns/ and /ʒ/ in /dʒ/ as in “gist” /ˈdʒɪst/.



The place of articulation of /tʃ/ and /dʒ/ is marked by an arrow (Kelly 2000, p.49)

Figure 6
The Illustration of the Place of Articulation of /tʃ/ and /dʒ/

In brief, the term “the place of articulation” of the English fricatives /ʃ/, /ʒ/ is a place in the mouth where the tongue makes light contact with the alveolar ridge and the airstream is partly constricted. As for the /t/ in /tʃ/ and /d/ in /dʒ/, the closure is made by the tongue and against alveolar ridge. So the airflow is completely stopped and then moves to the next fricative sounds.

The articulatory features help us to understand the speech sounds of /ʃ/, /ʒ/, /tʃ/ and /dʒ/ which are physically articulated. However, with the current development of high technology, we can understand the English sounds /ʃ/, /ʒ/, /tʃ/, /dʒ/ more properly and precisely.

1.3 Acoustic Features

“The acoustic structure of consonants is usually more complicated than that of vowels” because “the formant transitions of the same consonant are different for different vowels” (Ladefoged, 2000, p.179) and the consonants are dependent on the change of the vocal tract. Moreover, changing the shape of the vocal tract will change the formant frequencies and the length of consonants is very short in general and different kinds of sounds will have different visual representation of an acoustic signal in the spectrogram.

1.3.1 The English Fricatives

Fricatives consonants are “produced by forming a narrow constriction at some point along the length of the vocal cord tract and generating turbulence noise in the vicinity of this supraglottal constriction. The glottal opening is usually adjusted to be greater than the cross-sectional area of the supraglottal constriction, at least for the most common types of fricative consonants.” And there are two attributes which distinguish the production of fricative consonants from unaspirated stop consonants: (1) “the supraglottal articulators form a narrow constriction rather than a complete closure, and (2) some adjustment of the glottal opening is made to maximize the amplitude of the

turbulence noise that is generated near the supraglottal constriction” (Stevens, 2000, p.378).

1). The voicing of the English fricatives. The English sound /ʃ/ is voiceless and /ʒ/ is voiced. The voiced sound is “characterized by the supposition of harmonic sounds upon the noise source of the voiceless sound.” The spectrum of voiced consonants, the most striking manifestation of “voicing” is the appearance of a strong low component which is represented by the voice bar along the base line of the spectrogram.” In production “voiced phonemes are emitted with concomitant periodic vibrations of the vocal bands and voiceless phonemes without such vibration” (Jakobson, Fant and Halle, 1967, p.23). Moreover, the voiced fricatives have formant-like bands nearly the same as the voiceless counterpart. However, “they often show vertical striations, and sometimes a voice bar, a formant-like band of energy at very low frequencies” (Rogers, 2000, p.162) while the voice bar of voiceless fricatives is blurred or faint because the vocal cords do not vibrate during the sound production.

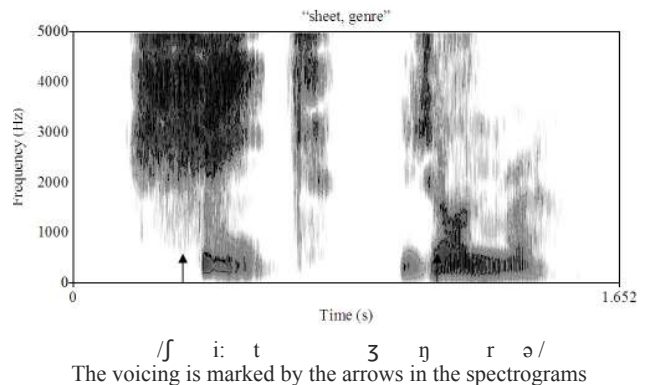


Figure 7
The Spectrograms of the Words “Sheet”, “Genre” as Spoken in an RP Accent by a Female

2). The manner of articulation of the English fricatives. Fricatives are formed by sending the airstream through “a narrow constriction which causes it to become turbulent, generating an aperiodic source of sound. The aperiodicity, or frication, is the defining manner cue to fricatives” (Pisoni and Remez, 2005, p.193).

First, the F1 frequency of a voiced fricative sound is also related to the manner of articulation. According to Stevens (2000, p.385), if the frequency of the first formant is low, the airflow escapes through a narrow channel of the articulators and if the frequency of the first formant is high, the airflow from the lungs through a wider channel of the articulators. In other words, the frequency of the first formant is dependent on the narrowing of the supraglottal constriction in the oral cavity.

The English /ʃ/ and /ʒ/ are fricatives. Therefore, they are “characterized by the random noise pattern” or friction which is considered to be as one more acoustic cue for

both the English voiceless fricative /ʃ/ and the English voiced fricative /ʒ/.

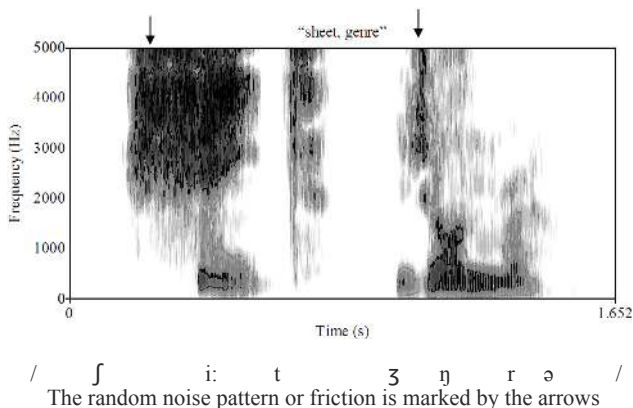


Figure 8
The Spectrograms of the Words “Sheet”, “Genre” as Spoken in an RP Accent by a Female

3). The place of articulation of the English fricatives. The acoustic cue to the place of articulation of the fricatives are based on the F2 frequency. The high frequency or low frequency depends on “the location and shape of constriction” (Stevens, 2000, p.411). To be easy to visualize, let’s look at the Figure 9.

This model is a general model of fricative production, whereby airflow becomes turbulent both in a construction and at an obstacle. The spectral shaping of turbulence is mostly performed by the front cavity, although there are minor effects caused by the back cavity. The presence of the obstacle in the front cavity also affects resonant characteristics.

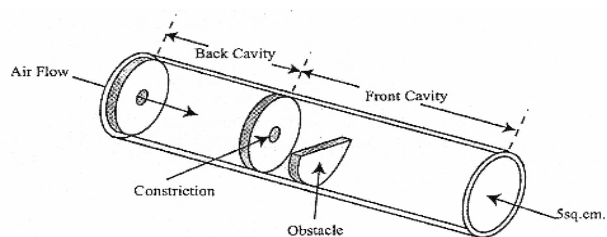


Figure 9
Physical Model of Fricative Production of Shadle’s (<http://www.pho.ucl/courses/spsci/b214/week2-6pdf>)

Stevens (2000, p.403) also describes the manner of articulation of the English fricatives /ʃ/, /ʒ/ that the point of maximum constriction is a few millimeters posterior to the alveolar ridge, and the tongue blade is shaped in such a way as to produce a relatively long and narrow channel behind the point of maximum constriction. This channel is formed between the surface of the tongue blade and the hard palate. The elevated tongue blade tends to create a space between the lower surface of the tongue blade and the floor of oral cavity. As for the alveolar

fricative consonants, the airstream emerging from the constriction is directed against the lower teeth, resulting in a turbulence noise source at this obstacle. For that reason, the remainder of the vocal tract posterior to the palatal constriction has a cross-sectional area that is large in comparison with the cross-sectional areas of the various components of the anterior part. In short, the acoustic cue for /ʃ/ and /ʒ/ is the F2 frequency.

1.3.2 The English Affricates

According to Pisoni and Remez (2005, p.195), the English affricates are sequences of the oral alveolar stops and the post-alveolar fricatives that listeners integrate perceptually into a single phonology entity. The sequence of acoustic segments is generated by the oral alveolar stop plus the post-alveolar fricative sequence. Therefore, the English affricates contain acoustic cues for both stops and fricatives. So the acoustic cues for the English affricates /tʃ/, /dʒ/ are more complicated than those of the English fricatives /ʃ/, /ʒ/.

1). The voicing of the English affricates. The English affricates are “produced by stopping the airstream from the lungs, and then slowly releasing it with a friction”. “The first part of an affricate is similar to a stop” and the second part is similar to a fricative” (Richards, Platt and Platt, 1992, p.11).

The voice bar of the affricate /tʃ/ is not evident or faint. Therefore, the affricate /tʃ/ is a voiceless sound. On the contrary, the voice bar of the affricate /dʒ/ is evident or clear. So the affricate /dʒ/ is voiced (see Figure 10).

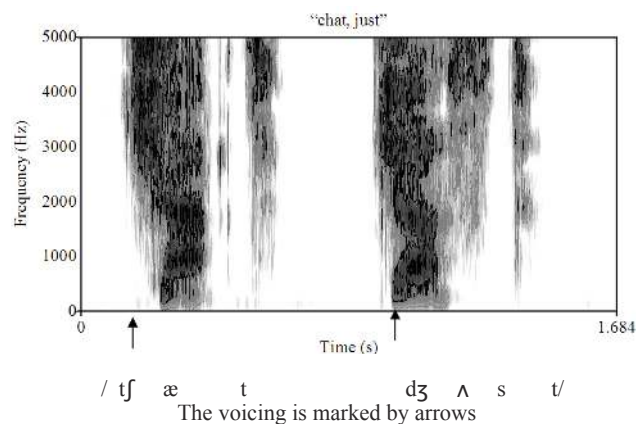


Figure 10
The Spectrograms of the Words “Chat”, “Just” as Spoken in an RP Accent by a Female

However, the voicing of the stops /t/ and /d/ in the affricate /dʒ/ and /tʃ/ are varied. According to Ogden (2009, p.99) there are at least “4 kinds of voicing of oral stops” such as fully voiced oral stops, partially voiced oral stops, voiceless aspirated oral stops, and voiceless preaspirated oral stops.

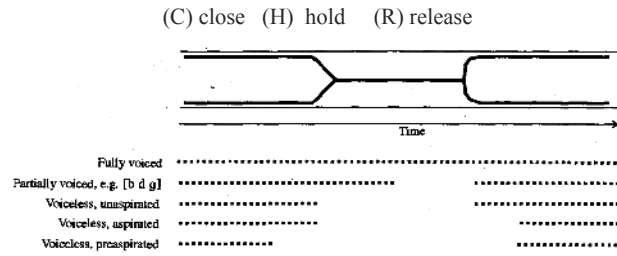


Figure 11
Voicing for Oral Stops. Dotted Line = Voicing. The Upper Lines Represent Two Articulators Moving Together and Apart (Ogden, 2009, p.100)

First, the voiced oral stops are fully voiced. “With a complete oral closure and eventually the pressure above and below the glottis equalizes, because the air above the glottis has no means of escape. Therefore, the duration of a fully voiced closure is constrained by how long a pressure difference across the glottis can be maintained. If the closure is released as soon as voicing stops, a fully voiced plosive is produced. Fully voiced plosives commonly occur between two voiced sounds as in “doodle”, “middle” (Ogden, 2009, p.100).

Second, the voiced oral stops are partly voiced. For example, the sound /d/ in English is generally partly voiced if it is preceded by a voiced sound and utterance-final /d/ is often partly voiced, and has no voicing, or whisper voice, on release. (Ogden, 2009, p.100).

Third, the voiceless stop /t/ is aspirated. The voicing stops at about the same time as the closure is made. Throughout the closure, the vocal fold vibration is held open, so the hold phrase is voiceless (Ogden, 2009, p.102).

Finally, the voiceless oral stop /t/ is preaspirated. “If voicing stops before a complete closure is achieved, i.e. the vocal folds allow air to pass through while the closure is still being made, this results in a short period of voicelessness and friction as the closure is being made, and is commonly known as preaspiration.” Because “the source of noise is turbulent at the glottis. This source of noise is transcribed as [th] and is known as preaspiration” (Ogden, 2009, p.104).

2). The manner and place of articulation of the English affricates. According to Stevens (2000, p.412), the English affricates are sequences of the English stops /t/, /d/ and the fricatives /ʃ/, /ʒ/. The manner and place of the second part of the English affricates are similar to fricatives /ʃ/, /ʒ/ which are presented above. Thus, the acoustic cues to the manner and place of the /t/ and /d/ in /tʃ/, /dʒ/ are presented as follows:

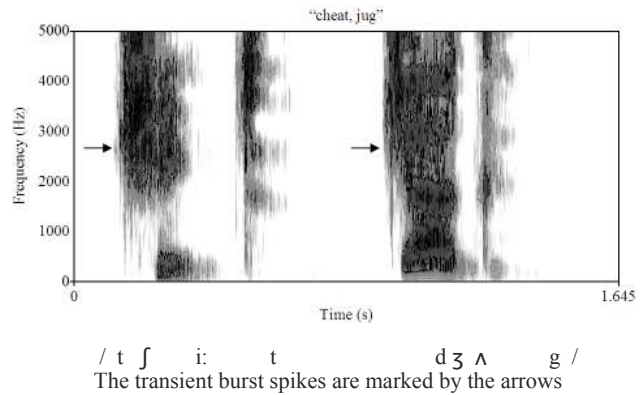


Figure 12
The Spectrogram of the Word “Cheat” and “Jug” as Spoken in an RP Accent by a Male

The acoustic cues to the manner of articulation of /t/ and /d/ are mainly based on “the vertical spike marking the release” (Pisoni and Remez, 2005, p.183). “The release of the aspirated stops is marked by a sharp spike corresponding to the onset of a burst of noise. This noise has a comparatively random pattern, mainly in the upper frequencies. (Ladefoged, 2000, p.180, 181). In addition, the acoustic cue to the manner of articulation is also based on “the relative silence of the closure interval” (Pisoni and Remez, 2005, p.183). Because the oral stops involve a total occlusion of the vocal tract, and thus a complete filter, i.e. no resonance being contributed by the vocal tract. The result of a period of silence in the spectrogram is known as a gap. This gap is used to distinguish the oral stops from fricatives.

As for the place of articulation of /t/ and /d/, the acoustic cues to the place of articulation of oral stop consonants are mainly based on the center burst frequency, the second formant frequency (Pisoni and Remez, 2005, p.183). In brief, the acoustic cues to the manner of /t/ in /tʃ/ and /d/ in /dʒ/ are the vertical spikes marking the release and the silent closure interval. The acoustic cues to the place of articulation are based on center burst frequency and the F2 frequency.

1.4 The Consonant Clusters

1.4.1 Definition of Consonant Clusters

A consonant cluster is “a sequence of two or more consonants.” (Richards, Platt and Platt, 1992, p.79) Consonant clusters in English may occur in word-initial position as in “draw”, in word-medial position as in “instrument” and in word-final position as in “garaged”. According to Roach (1990, p.68), “when we have two or more consonants together we call them a consonant cluster.” The number of consonants which go together is varied. “The largest initial consonant cluster in English contains three consonants” as in “strong” (Kuiper and Allan, 1996, p.64). In contrast, there are no consonant clusters in modern Vietnamese.

1.4.2 Common Consonant Clusters go with /ʃ/, /ʒ/, /tʃ/ and /dʒ/

English has a “limited number of permitted combinations of consonants in initial and final clusters.” (Avery and Ehlich, 1995, p.54) Some of the following consonant clusters are taken from Avery and Ehlich (1995, p.55) such as /-ʃl/, /-lʃ/, /-rʃ/, /-ltʃt/, /-rtʃt/, and /-ndʃd/.

Table 1
Consonant Clusters Consist of /ʃ/

No.	Distribution	Consonant clusters with /ʃ/	Examples
1	In word-initial position	/ʃr/	shrink, shrine, shroud
2		/-ʃl/	crucial, essential
3		/-lʃ/	Welsh
4	In word-final position	/-ʃn/	mention, nation
5		/-ʃt/	washed, pushed
6		/-rʃ/	marshed
7		/-ʃnz/	functions, nations
8		/ʃt-/	shtup, shtum, shtetl

Table 2
Consonant Clusters Consist of /ʒ/

No.	Distribution	Consonant clusters with /ʒ/	Examples
1	In word-final position	/-ʒd/	garaged, sabotaged
2		/-ʒn/	vision, division

Table 3
Consonant Clusters Consist of /tʃ/

No.	Distribution	Consonant clusters with /tʃ/	Examples
1	In word-final position	/-ltʃt/	belched
2		/-tʃt/	hitched, watched
3		/-rtʃt/	arched

Table 4
Consonant Clusters Consist of /dʒ/

No.	Distribution	Consonant clusters with /dʒ/	Examples
1	In word-final position	/-ldʒd/	bulged,
2		/-ndʒd/	changed, exchanged

1.5 Mother Tongue's Interference in Language Acquisition

1.5.1 The Theoretical Background

Second language acquisition is a complex process involving many interrelated factors. According to Richards, Platt and Platt (1992, p.325), “second language acquisition is the process by which people develop proficiency in a second language or foreign language. These processes are often investigated with the expectation that information about them may be useful in language teaching.” The term “second language acquisition” includes “longitudinal studies and cases studies of the

development of syntax and phonology in second and foreign language”, “analysis of spoken and written discourse of second and foreign language learners”, and the study of other aspects of language development.” From another point of view, “second language acquisition refers to the subconscious or conscious process by which a language other than the mother tongue is learnt in a natural or tutored setting” (Ellis, 1985, p.6). Failing to overcome the L1 patterns of phonology in speaking the second language will result in foreign accents. In other words, foreign language learners are much affected by the phonological patterns of their mother tongue.

“Each language has a different inventory of sounds, different rules for combining these sounds into words, and different stress and intonation patterns.” The role of the native language influence foreign language learners in three ways (Avery and Elrich, 1995, p.xv):

First, “difficulties may arise when a learners encounters in English that are not apart of the sounds inventory of the learner’s native language.” Learners have difficulty in exercising their mouth in the particular way required to pronounce the English sounds correctly. For example, Vietnamese learners find it hard to pronounce the English sounds correctly: /p/, /ð/, /θ/, /ʃ/, /ʒ/, / tʃ/, /dʒ/, etc. Let’s take the English sound /p/ as an example, / p/ is an inventory sounds which does not really exist in Vietnamese except in writing. The English /p/ can occur in word-initial, word-medial and word-final position. For that reason, Vietnamese learners tend to substitute the Vietnamese sound /b/ for the English /p/. Not only that, the English /p/ is aspirated if it occurs in word-initial position and in a stressed syllable as in “pen”.

Second, “difficulties may arise because the rules for combining sounds into words are different in the learner’s language. This type of difficulty can occur even when a particular sound is a part of the inventory of both English and the native language.” For example, the combination of Vietnamese words can be a vowel + a vowel as in “ai”, consonant + (a) vowel(s) as in “hai”, (a) vowel (s) + a consonant as in “ít” and a consonant + (a) vowel (s) + a consonant as in “khuong” or “nam.” In addition, there are no consonant clusters in modern Vietnamese and the final sounds in Vietnamese are never released. On the contrary, certain English consonants go together to form consonant clusters as in “stray” or watched”. The combination of English words is much more complicated than those in Vietnamese because English is a polysyllabic language and Vietnamese is a monosyllabic one. English has the same characteristics as those of Vietnamese, the sound contours of words are longer as in “ international” (a vowel +consonant clusters + vowel + a consonant + a vowel + consonant clusters + a vowel + a consonant). Besides that, linking and assimilation are also “common features of speech” in English.

Third, “the pattern of stress and intonation, which determine the overall rhythm and melody of a language,

can be transferred from the native language into the second language”. For example: “Ban là thư ký mới không? (Are you a new secretary?) The Vietnamese use the falling intonation. The English people use three kinds of tones with different meanings. If the speaker uses the falling tone, he or she just want to make sure that than to seem to be finding out. If the speaker uses the rise tone, he or she wants to find out who you are? If the speaker uses the fall – rise tone, the sentence means like “am I right in thinking you are the new secretary (the person I have heard so much about) (Brazil, 2006, p.62).

In general, “the native language not only affects the ability to produce English sounds but also the ability to hear English sounds experienced teachers certainly know the frustration involved in having students continually repeat a mispronounced word in the same way. Students may seem impervious to correction but, in fact, the problem often arises because the world is heard through the sound system of the native language. Thus, sounds which occur in the native language will be heard rather than the actual sounds of English which are being produced by the teacher. This highlights a very important point concerning the influence of the native language. It is if learners hear the second language through a ‘filter’, the filter being the sound system of the native language” (Avery and Ehrlich, 1995, p.xv).

In summary, the second language learners’ errors may be the predictable and potential difficulties that the student subjects face because of the influences of Vietnamese.

1.5.2 The Vietnamese Consonants Affect the Students’ Acquisition of the English /f/, /z/, /tʃ/ and /dʒ/

All of the teachers who participate in this research confirm that the third-year in-service training students who major in English have difficulty in pronouncing a number of certain English consonants. The English consonants that the students have the most problem with are the English fricatives /f/, /z/ and the English /tʃ/, /dʒ/. Their difficulties are caused by the following reasons:

First, Vietnamese is a tone, monosyllabic language, whereas the English is a polysyllabic one. According to Phan (2007, p.3), there are 21 consonant phonemes in Vietnamese. Most of the Vietnamese consonants occur in word-initial position. Only 8 of them can occur in word-final position while the English language has 24 consonants and most of them can occur in word-initial position, word-medial position and word-final position except for the English consonant /h/ which does not occur in word-initial position and the English consonant /z/ can occur in word-initial position, but the number of words are not many. For these above reasons, the student subjects have the most difficulty in pronouncing the English consonants in word-medial and word-final position.

Second, the students also tend to forget to pronounce and release the English consonants in word-final position. According to Avery and Ehrlich (1995, p.154) fricatives do not occur in word-final position in Vietnamese.

Therefore, Vietnamese speakers may omit fricatives at the end of words. In addition, Vietnamese speakers pronounce the consonants in word-final position in a state but do not release them in the same way as in English. Frankly speaking, this fault is partly caused by the influence of Vietnamese.

Finally, some characteristics of the English sounds seem have something in common with the Vietnamese ones. However, they differ internally from each other. For example, the English /tʃ/ as in “church” and the Vietnamese /c/ as in “cho” (to give) are voiceless and laminal (tongue blade). The English /tʃ/ is post-alveolar affricate (Roach, 1990, p.62) whereas the Vietnamese /c/ is palatal stop. Therefore, they cannot be counterparts or equivalents.

2. METHODOLOGY

2.1 Participants

The study was implemented with the involvement of 85 third-year students from Hung Vuong University. Here are the descriptions of the participants collected through the questionnaire survey.

Question 1 investigates student subjects’ age.

Table 5
The Students’ Age

Age	Responses	
	N	Percentage (%)
Over 18 – 25	49	57.6%
26 – 30	27	31.8%
31– 35	8	9.5%
40 – 45	1	1.2%
Total	85	100.0%

The age range of the students is from 18 to 25 years old (57 %); from 26 to 30 makes up 31%; from 31 to 35 is 9.5% and from 40 to 45 is only 1.2 %. Table 5 shows that majority of the students are under 30 years old.

Question 2 surveys the students’ gender.

Table 6
The Students’ Gender

Gender	Responses	
	N	Percentage (%)
Male	19	22.4%
Female	66	77.6%
Total	85	100.0%

Majority of the students are female (77.6 %, 66 persons) and the rest of the student subjects are male (22.4%, 19 persons). Table 6 shows the distribution of the students by gender.

Question 3 asks students about how long they have

been learning English.

Table 7
The Length of English Learning (Year)

The length of English learning	Responses	
	N	Percentage (%)
5 – 7	26	30.6%
8 – 10	45	52.9%
Over 10	14	16.5%
Total	85	100.0%

As Table 7 shows, more than half of the students have been learning English from 8 to 10 years (52.9%, 45 persons). Learning English from 5 to 7 years constitutes 30.6% (26 persons) and the students studying English over 10 years are only 16.5%.

2.2 Instrument

The questionnaire was used for the student subjects. With 13 questions, the student s' questionnaire is divided into two parts. The first part (Questions 1 to 3) collects the students' personal information. The second part (Questions 4 to 13) aims at gathering data on aspects related to students' acquisition of the English consonants in general and the English fricatives /f/, /z/ and the affricates /tʃ/, /dʒ/ in particular. The questionnaire for students was written in Vietnamese so that students can understand the questions clearly and accurately express their points of view.

3. FINDINGS FROM THE QUESTIONNAIRE SURVEY

Question 4 investigates the students' regional accents in Vietnam.

Table 8
Students' Regional Accents in Vietnam

Accents	Responses	
	N	Percentage (%)
Northern accent	10	11.8%
North central coast accent	5	5.9%
South central coast and southern accent	70	82.4%
Total	85	100.0%

As illustrated by Table 8, 70% of the student subjects speak Vietnamese with the south central coast and southern accent. The students with northern accent are only 11.8 % and the rest of the students with north central coast accent are only 5%.

Question 5 investigates the student subjects whether they have problems with a certain number of the English consonants or not.

Table 9
Students Have the English Consonant Problems

Students have the English consonant problems	Responses	
	N	Percentage (%)
Totally agree	63	74.1%
Partially agree	16	18.8%
Totally disagree	2	2.4%
Partially disagree	3	3.5%
Neither agree or disagree	1	1.2%
Total	85	100.0%

Table 9 indicates that 74.1% of the student subjects agree completely that they have difficulty with a number of the English consonants; 18.8% of all the student subjects agree partially; 2.4% of the students disagree totally; 3.5% of the students disagree partially and 1.2% of the student subjects agree or disagree. In short, majority of the students have difficulty with a number of the English consonants.

Question 6 asks the student subjects which English consonants that they have difficulty with.

Table 10
Students' Difficult English Consonants

English consonants	Responses	
	N	Percentage (%)
/p/	19	4.1%
/b/	15	3.2%
/t/	7	1.5%
/d/	7	1.5%
/ŋ/	12	4.5%
/f/	4	0.9%
/v/	1	0.2%
/ð/	43	9.3%
/θ/	34	7.4%
/s/	10	2.2%
/z/	15	3.2%
/ʃ/	54	11.7%
/ʒ/	67	14.7%
/h/	5	1.1%
/tʃ/	67	14.5%
/dʒ/	80	17.3%
/w/	4	0.9%
/r/	4	0.9%
/j/	5	1.1%
Total	462	100.0%

Table 10 illustrates that the student subjects have difficulty with 19 English consonants; the bilabial plosives /p/ and /b/ constitute 4.1% and 3.2%; the alveolar plosives /t/ and /d/ are 1.5% and 1.5%; the velar nasal /ŋ/ is 4.5%; the labio-dental fricatives /f/ and /v/ are 0.9% and 0.2%; the dental fricatives /ð/ and /θ/ are 9.3% and 7.4%; the alveolar fricatives /s/ and /z/ are 2.2% and 3.2%; the post-alveolar fricatives /ʃ/ and /ʒ/ are 11.7% and 14.7%; the glottal fricative /h/ is only 1.1%; the post-alveolar affricates /tʃ/ and /dʒ/ make up 14.5% and 17.3%; the bilabial approximant is 0.9%; the post-alveolar /r/ is also 0.9% and the palatal approximant /j/ constitutes 1.1%.

Question 7 investigates whether the student subjects

have the most difficulty with the four English sounds /f/, /z/, /tʃ/ and /dʒ/ or not.

Table 11
Students' Problematic Consonants

Students' problematic consonants	Responses	
	N	Percentage (%)
Totally agree	52	61.2%
Partially agree	30	35.2%
Totally disagree	1	1.2%
Partially disagree	0	0.0%
Neither agree or disagree	2	2.4%
Total	85	100.0%

According to the survey of the four English consonants /f/, /z/, /tʃ/ and /dʒ/, 61.2% of the student subjects agree totally. 35% of the student subjects agree partially. Only 1.2% of the student subjects disagree totally and 2.4% of the student subjects neither agree nor disagree. So the majority of the student subjects have difficulty with these English consonants.

Question 8 surveys the student subjects about the positions of /f/, /z/, /tʃ/ and /dʒ/ that they have difficulty with.

Table 12.
The Rate of the Difficult Positions of /f/

/f/	Responses	
	N	Percentage (%)
In word-initial position	27	25.7%
In word-medial position	40	38.1%
In word-final position	38	36.2%
Total	105	100.0%

As for the post-alveolar fricative /f/, Table 12 shows that 25.7% of the student subjects have difficulty with /f/ in word-initial position. 38.1% of the student subjects have trouble with /f/ in word-medial position and 36.2% of the student subjects find the English consonant /f/ in word-final position hard. In short, the most difficult positions that the student subjects face are the medial position and the final position.

Table 13
The Rate of the Difficult Positions of /z/

/z/	Responses	
	N	Percentage (%)
In word-initial position	34	24.8%
In word-medial position	45	32.8%
In word-final position	58	42.3%
Total	137	100.0%

From Table 13, 24.8% of the student subjects have difficulty with /z/ in word-initial position. 32.8% of the student subjects have trouble with /z/ in word-medial

position and 42.3% of the student subjects find the sound /z/ in word-final position the most difficult.

Table 14
The Rate of the Difficult Positions of /tʃ/

/tʃ/	Responses	
	N	Percentage (%)
In word-initial position	34	27.0%
In word-medial position	38	30.2%
In word-final position	54	42.9%
Total	126	100.0%

As Table 14 illustrates that 27% of the student subjects find /tʃ/ in word-initial position hard; 30% of the student subjects have considerable difficulty with /tʃ/ in word-medial position and 42% of the student subjects have difficulty with /tʃ/ in word-final position.

Table 15
The Rate of the Difficult Positions of /dʒ/

/dʒ/	Responses	
	N	Percentage (%)
In word-initial position	64	21.5%
In word-medial position	65	32.0%
In word-final position	74	36.5%
Total	203	100.0%

As for the English post-alveolar affricate /dʒ/, the student subjects have difficulty with all of the three positions in which the English affricate /dʒ/ can occur. The English /dʒ/ in word-initial position is 31.5%; the English /dʒ/ in word-medial position makes up 32% and the English /dʒ/ in word-final position constitutes up to 36.5%.

Question 9 collects students' opinion about the sound practice methods that they like best.

Table 16
The Sound Practice Methods Students Like Best

The sound practice methods students like best	Responses	
	N	Percentage (%)
Practice in discrete sounds	82	20.6%
Practice in individual words	81	20.3%
Practice in word phrases	80	20.1%
Practice in individual sentence	77	19.3%
Practice in sentence contexts	79	19.8%
Total	399	100.0%

Table 16 reveals that 20% of the student subjects like practicing the English consonants in discrete sounds. 20.3% of the student subjects love practicing the English consonants in individual words. 20.1% of the students are keen on practicing the English consonants in word phrases. 19.3% of the students are interested in practicing the English consonants in individual sentence and 19.8% of the students are fond of practicing the English sounds

in sentence contexts. Here are the orders that should be used in the experimental teaching.

1. Practice in discrete sounds
2. Practice in individual words
3. Practice in word phrases
4. Practice in sentence contact
5. Practice in individual sentence

Question 10 invites the students' opinion about which activities make the class interesting and successful.

Table 17
The Activities Make the Class Interesting

The activities make the class interesting	Responses	
	N	Percentage (%)
Repeat after the recordings	83	25.3%
Repeat after the teacher	86	26.2%
Play games	78	23.8%
Listen to English songs	81	24.7%
Total	328	100.0%

As can be seen from Table 17, 25.3% of the student subjects like repeat after the recordings. 26.2% of the student subjects love repeating after the teacher. 23.8% of the student subjects are interested in playing games and 24.7% of the student subjects are keen on listening to English songs. Here is the important order of the activities used in the experimental teaching.

1. Repeat after the teacher
2. Repeat after the recordings
3. Listen to English songs
4. Play games

Question 11 surveys the student subjects about whether they have difficulty with consonant clusters that contain /f/, /z/, /tʃ/ and /dʒ/ or not.

Table 18
The Students Have Difficulty with the Consonant Clusters

The students have difficulty with the consonant clusters	Responses	
	N	Percentage (%)
Totally agree	69	81.2%
Partially agree	15	17.6%
Neither agree or disagree	1	1.2%
Total	85	100.0%

As Table 18 reveals that 81% of the student subjects agree totally; 17.6% of the student subjects agree partially and only 1.2% of the student subjects neither agree or disagree. It is concluded that the majority of the student subjects have trouble with the English consonant clusters consisting of /f/, /z/, /tʃ/ and /dʒ/.

Question 12 asks the student subjects whether the correct pronunciation of the English /f/, /z/, /tʃ/ and /dʒ/ is important to their communication skill in English or not.

Table 19
The Impacts of Mispronunciation

The impacts of mispronunciation	Responses	
	N	Percentage (%)
Yes	82	96.5%
No	3	3.5%
Total	85	100.0%

The majority of the student subjects agree (96.5%) and only 3.5% of the student subjects do not think so. They also say that the mispronunciation of the four English sounds such as /f/, /z/, /tʃ/ and /dʒ/ can make the interlocutors misunderstood and confused. Moreover, the mispronunciation can make the conversation difficult and less interesting.

Question 13 aims at finding out what they have done to overcome their difficulty in pronouncing the English words, phrases and sentences containing /f/, /z/, /tʃ/ and /dʒ/.

Table 20
The Students' Ways to Overcome Their Difficulty

The students' ways to overcome their difficulty	Responses	
	N	Percentage (%)
Listen to the tapes, watch TV, films to imitate the sounds	40	39.6%
Imitate the teacher	23	22.8%
Have conversation with native speakers	10	9.9%
Omit the consonants	17	16.8%
Leave unanswered	11	10.9%
Total	101	100.0%

It is clear from Table 20 that 39.6% of the student subjects listen to tapes, watch TV and films to improve their English consonant learning. 22.8% of the students try to imitate their teachers' accent. 9.9% of the students sometimes have chances to practice English with native speakers. 16.8% of the students omit the learning the English consonants and 10.9% leave unanswered.

CONCLUSION

According to results of the questionnaire survey, the students have the most difficulty in pronouncing the English fricatives /f/, /z/ and the affricates /tʃ/ and /dʒ/ among the English consonants surveyed. The students tend to substitute Vietnamese sounds for the English sounds.

Firstly, students do not find it much difficult to pronounce the English /f/ in word-initial position as in "sheet". But, they have trouble pronouncing the English /f/ in word-medial position and word-final position as in "fishing" and "wash". They tend to substitute the Vietnamese /s/ as in "xua" (meaning "old") or /ʃ/ as in "sũa" (meaning "milk") for the English /f/ and they often

to forget to pronounce and release the English /f/ in word-final position. According to Avery and Ehrlich (1995, p.154), “Vietnamese speakers may omit fricatives at the ends of words”.

Secondly, the students have pronunciation problem with /ʒ/ in word-initial position, word-medial position and word-final position. They tend to substitute the Vietnamese /ʒ/ as in “di” (aunt) and “giò” (hour) for the English /ʒ/. They usually forget to pronounce and release the English /ʒ/ in word-final position and in English consonant clusters.

Thirdly, as for the English voiceless affricate /tʃ/ which is one of the most difficulties for the student subjects, they often substitute the Vietnamese stop /c/ as in “cho” (to give) for the English affricate /tʃ/ because they may not realize that the affricate /tʃ/ is the sequence of the English stop /t/ and the English fricative /ʃ/, whereas the Vietnamese consonant /c/ is a voiceless palatal stop. In addition, the student subjects cannot articulate the aspirated /t/ in /tʃ/ properly because the consonant /t/ in /tʃ/ is aspirated when it is in word-initial position and in a stressed syllable.

Finally, for the English voiced affricate /dʒ/, the student subjects with northern accent, north central coastal accent and south central coast sometimes substitute the Vietnamese /c/ as in “chị” (sister) for the English /dʒ/. The student subjects with south central coast and southern accent substitute the Vietnamese /c/ as in “chị” (sister) (Vuong, 1998, p.23) for the English /dʒ/ in word-initial position, in-word-medial position and in word-final position.

REFERENCES

- Avery, p.& Ehrlich, S. (1995). *Teaching American English Pronunciation*. Hong Kong: Oxford University Press.
- Brazil, D. (2000). *Pronunciation for Advanced Learners of English*. Cambridge: Cambridge University Press.
- Celce-Maurcia, M., Brinton. M.D., & Woodwin, M.J. (2002). *Teaching Pronunciation*. The UK: Cambridge University Press.
- Ellis, R. (1985). *Understanding Foreign Language Acquisition*. New York: Oxford University Press.
- Fromkin, V., Rodman, R., Collins, P., & Blair, D. (1988). *An Introduction to Language*. Australia: Harcourt Brace Jovanovich Group.
- Jakobson, R., Fant, M.G.C., & Halle, M. (1967). *Preliminaries to Speech Analysis*. The United States of America: The M-I-T Press.
- Kelly, G. (2000). *How to Teach Pronunciation*. Malaysia: Pearson.
- Kuiper, K. & Allan, S.W. (1996). *An Introduction To English Language Sound, Word and Sentence*. Hong Kong: Macmillan.
- Ladefoge, p.(2000). *A Course in Phonetics*. Thomson Wadsworth.
- Ladefoged, p.& Maddieson, I. (1996). *The Sounds of the World's Languages*. The United States of America: Blackwell Publishers.
- Ogden, R. (2009). *An Introduction to English Phonetics*. Great Britain: Edinburgh University Press.
- Pisoni, B.D. & Remez, E.R (2005). *The Handbook of Speech Perception*. Hong Kong: Blackwell Publishing.
- Phan, V.G. (2007). *Vietnamese for Beginners - Book 1*. Ho Chi Minh City: Tre Publishing House.
- Richards, C.J, Platt, J. & Platt, H. (1992). *Dictionary of Language Teaching & Applied Linguistics*. Singapore: Longman Publishing House.
- Rogers, H. (2000). *The Sounds of Language - An Introduction to Phonetics*. Singapore: Pearson.
- Roach, p.(1990). *English Phonetics and Phonology - A Practical Course Section*. London: Cambridge University Press.
- Steven, N.K. (2000). *Acoustic Phonetics*. The United of America: The MIT Press.
- Vuong, H.L. (1998). *Ngữ Âm Tiếng Việt Tin Giản*. Thua Thien Hue: Giao Duc Publishing House.
- Yates, J. (2003). *Pronounce it Perfectly in English*. Hauppauge, NY: Barron's Educational Series.