Mozaik Humaniora
Vol. 24 (2): 180-193

DOI 10.20473/mozaik.v24i2.56495

English Voiceless Stop Production of Javanese University Students in Yogyakarta

(Produksi Konsonan Hambat Letup Bahasa Inggris oleh Mahasiswa Berbahasa Jawa di Yogyakarta)

Prahastuti Nastiti Hadari[1*]
Tofan Dwi Hardjanto[2]
[1]Language Training Center, Universitas Muhammadiyah Yogyakarta
[1]Brawijaya Street, Bantul, Yogyakarta 55183
[1]Tel.: +6281227404760
Department of Language and Literature, Universitas Gadjah Mada
Nusantara Street No.1, Bulaksumur, Yogyakarta 55281
Tel. : +62(274)588688
[*]Email: prahastutinh16@gmail.com

Received: 16 April 2024 Revised: 12 December 2024 Published: 31 December 2024

Abstrak

Sebagai bahasa yang dipelajari hampir di seluruh dunia, variasi bahasa Inggris bermunculan dengan membawa karakteristik bahasa ibu pembelajar asing. Berdasarkan asumsi tersebut, penelitian ini bertujuan untuk mengidentifikasi pengaruh elemen fonologis bahasa Jawa terhadap produksi konsonan hambat letup bahasa Inggris oleh mahasiswa berbahasa Jawa berdasarkan gender. Data berupa rekaman suara dikumpulkan dari 36 mahasiswa yang terdiri atas 18 laki-laki dan 18 perempuan dari dua universitas terkemuka di Yogyakarta. Dengan menggunakan Praat, *voice onset time* (VOT) setiap bunyi dicatat dan dianalisis menggunakan Microsoft Excel. Hasil analisis menunjukkan bahwa berdasarkan rata-rata VOT sebanyak 514 data, baik siswa laki-laki maupun perempuan menghasilkan /p/ dengan hambatan pendek dan /t/ dan /k/ dengan hambatan panjang, sehingga membentuk pola /k/ > /t/ > /p/. Konsonan hambat letup tak bersuara pada posisi awal memiliki VOT yang lebih panjang dibandingkan posisi tengah. Mahasiswa laki-laki lebih banyak menghasilkan pergeseran fonem dibandingkan mahasiswa perempuan. Fonologi bahasa Jawa lebih banyak memengaruhi mahasiswa laki-laki dalam memproduksi [ʈ], yakni bentuk retrofleks dari *t-flap* [ɾ] sebagai varian /t/ bahasa Inggris pada posisi tengah. Namun, mereka menghasilkan VOT yang lebih panjang dibandingkan mahasiswa perempuan.

Kata kunci: fonetik, fonologi, konsonan hambat letup, *voice onset time*

Abstract

As a language taught almost everywhere in the world, English variations emerge, and they bring the characteristics of foreign learners' first languages. Under this assumption, this study aimed to seek how Javanese phonological elements interfere with the production of English voiceless

stops by Javanese students on a gender basis. Data in the form of sound recordings were collected from 36 students which consisted of 18 males and 18 females from two prominent universities in Yogyakarta. Using Praat, the voice onset time (VOT) of each sound was noted and processed using Microsoft Excel. The study result showed that based on the VOT means of 514 data, regardless of the position, both male and female students produced short-lagged /p/ and long-lagged /t/ and /k/, forming a pattern of /k/ > /t/ > /p/. Voiceless stops at the initial position have longer VOT than the medial position. Male students produced more phoneme shifts than female students. Male students seemed to be more Javanese-influenced in the way they produced more [ʈ], the retroflex version of t-flap [ɾ] as the English /t/ variation, at the medial position although they generally produced longer VOT than female students.

Keywords: phonetics, phonology, stop consonants, voice onset time

INTRODUCTION

English is a language used and taught in every region across the globe. As a means of communication, a language is mainly spoken. That is when cross-linguistic transfers occur because not all English phonemes are used in a foreign learner's first language. It is only reasonable that elements in the first language are brought into their English pronunciation (Genetti 2014), creating new variations of English (Bolton 2006; Nelson et al. 2020; Kachru et al. 2006; Kirkpatrick 2010).

In Indonesia, most people speak more than one language. Due to the multilingual situation where the country has more than 300 ethnicities (Suku Bangsa 2017), it is possible that one person speaks three languages: a national language, a vernacular language, and an international language. Javanese is a vernacular language spoken by more than 95 million people, which is at least 42% of the total population (Purbowati et al. 2024; Statistics Indonesia 2015), and English is a foreign language that is mostly learned by Indonesian students. In this situation, cross-linguistic transfers from Javanese to English can possibly occur.

English and Javanese have different phonological systems and rules, including the stops. Stops are consonants that are produced with total airflow obstruction. In English, the most common consonants regarded as voiceless stops are /p/, /t/, /k/, which are commonly aspirated at the beginning of a stressed syllable, usually in one-syllable words and almost unaspirated at the beginning of an unstressed syllable, usually in more than one-syllable words (Carley & Mees 2020; Collins et al. 2019; Fromkin et al. 2014; O'Grady & Archibald 2016; Zsiga 2013). Javanese also has the three voiceless stops, also the retroflex /ʈ/, with opposite aspiration rules, where they are unaspirated regardless of the position and the syllable number (Marsono 2018; Marsono 2019; Thurgood 2004; Wedhawati et al. 2001).

With the possibility of language interference from Javanese into English, an earlier study by Mesthrie (2004, 1106) found that English /p/, /t/, and /k/ are usually unaspirated, and /t/ was produced as retroflex /ʈ/ by Indonesians, but it was not clear where the study was conducted. Several previous studies saw cross-linguistic transfers as errors (Senowarsito & Ardini 2019; Utami et al. 2017). Other than Mesthrie's finding, Ambalegin & Hulu (2019) sought further explanations on how transfers occur and found that students in Batam pronounced /t/ as unaspirated.

A commonly used measurement of voiceless stops, especially in English, is voice onset time, or VOT. VOT is the relation between the stop release and the beginning of a voicing

(Gussenhoven & Jacobs 2017). With the stop release regarded as the 0 point, when the voicing starts after the release, a voicing lag occurs, which is what happens to aspirated voiceless stops. Measured in milliseconds (ms), VOT ≥ 35 ms is considered long-lagged, VOT < 35 ms and ≥ 0 ms is considered short-lagged, both with a positive value, and VOT < 0 has a negative value (Chodroff et al. 2019). English aspirated voiceless stops are typically around 50 ms long, with /k/ at around 50–60 ms or longer and /p/ and /t/ less than that (Gussenhoven & Jacobs 2017; Ladefoged & Disner 2012).

Regarding VOT, a study found that 7–8-year-old children in Malang learning English extensively since two years old produced /p/ and /t/ at around 69 ms and /k/ at around 106 ms (Zen 2018). In Surabayan Javanese, it was found that /t/ had the longest duration, followed by /p/, and then /k/ (Soetjahyati 2018). VOT was found to correlate with gender where males produce relatively longer VOT (Oh 2019) and more variations in English (Yu et al. 2014), and /t/ became the most gender-related in Turkish (Korkmaz & Boyaci 2018).

Since Indonesia has many ethnicities with different vernacular languages, studies with a clear limitation on the language use in Indonesia are needed. Complementing the previous studies with the same topic, this study aimed at revealing how English voiceless stops spoken by Javanese university students in Yogyakarta are produced using VOT measurement results on a gender basis.

METHOD

Data in the form of sounds were collected online at the beginning of 2022 under the Covid-19 restrictions by conducting several steps. A screening test was done using Google Forms to find appropriate respondents in four English study programs at two universities in Yogyakarta. Those who passed the screening test were active students speaking Javanese as their first language and had English capability of at least B1. Next, they self-recorded their voices pronouncing a list of 15 isolated words with voiceless stops /p/, /t/, and /k/, 9 at the initial and 6 at the medial positions with different syllable numbers. From 18 male and 18 female students in the English Literature and the English Education Study Programs of Universitas Negeri Yogyakarta and the Applied English and the English Study Programs of Universitas Gadjah Mada, 540 data were collected using the same recording application called Voice Recorder at 44 KHz and mono 11-bit (Buchstaller & Khattab 2013) by each respondent at home.

The data were put into Praat and manually analyzed based on the spectrograms and sounds to get the VOT values. The measurement results were noted and processed using Microsoft Excel. VOT means and ranges were then put on tables and some of the spectrograms were screen-captured to be presented and discussed. Voiceless stops realized as other than /p/, /t/, /ʈ/, and /k/ and their variations were excluded from the VOT discussion. Instead, the phoneme shifts were noted to later be discussed whether there are differences in frequencies between male and female students.

RESULTS AND DISCUSSION

VOTs of Javanese Students' English Voiceless Stops

From a total of 540 data collected, the VOTs of 514 data were measured. The rest of the 26 data were not measured as they were not realized as either /p/, /t/, /ʈ/, or /k/. The VOT measurement analysis found that positions affect the stop consonant VOTs of the Javanese

students. Regardless of the position, this study found that the voiceless stop VOTs are all positive.

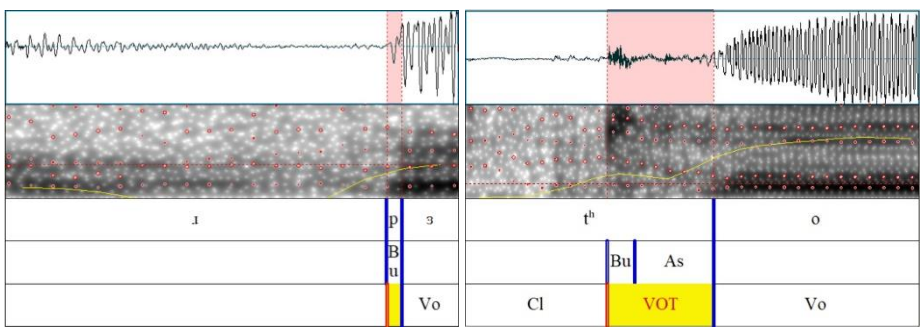
In general, /p/ has the shortest VOT duration among the three voiceless stops at 29 ms, followed by /t/ at 47 ms, and then /k/ at 60 ms. Therefore, /p/ is considered short-lagged, and /t/ and /k/ are considered long-lagged. This finding, especially for /t/ and /k/, matches the theory that English voiceless stops usually have around 50 ms VOT with /k/ around 60 ms and /t/ less than that (Gussenhoven & Jacobs 2017; Ladefoged & Disner 2012)

Table 1. VOTs of Javanese Students' English Voiceless Stops

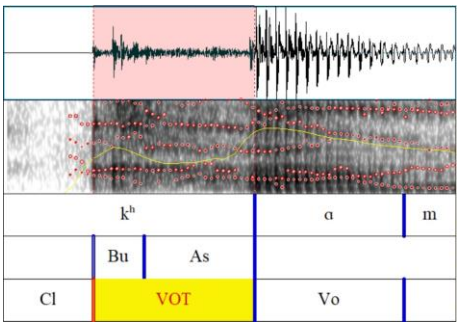
Sound	F			M		
	n	$\bar{x}$	Range	n	$\bar{x}$	Range
p	90	28	4–105	90	31	7–117
t	80	42	7–149	74	53	8–159
k	90	58	17–149	90	61	23–171

Ignoring the position, the students' voiceless stop production has a constant pattern applied for both male and female students. As can be concluded from Table 1, /k/ is the longest, followed by /t/, and then /p/ (/k/ > /t/ > /p/). This pattern is in line with the previous studies (Chodroff et al. 2019; Gussenhoven & Jacobs 2017; Ladefoged & Disner 2012; Soetjahyati 2018; Zen 2018). Moreover, /p/ is short-lagged, and /t/ and /k/ are long-lagged for both male and female students. The fact that /p/ is short-lagged does not mean that it is unaspirated because the values of both male and female VOTs are almost 35 ms. It means that all the voiceless stops are averagely aspirated, which is not in line with some previous studies (Ambalegin & Hulu 2019; Mesthrie 2004).

When compared, male students have longer VOT, although not very significant (Oh 2019), especially for /p/ and /k/. There is a quite bigger gap for /t/, but it is possibly related to male students' data that is smaller in number due to phoneme shifts, causing the data to not be proportional between the two gender categories. The ranges also show the same tendency that male students have longer VOT because they have higher minimum and maximum VOTs than female students. In each gender category, the range gaps for each stop consonant are not very much different between one and another but male students have relatively bigger ranges than female students. It means that male students have a slightly longer VOT for each voiceless stop but also bigger ranges.



(a) (b)



(c)

Figure 1. Production of (a) /p/ in 'composer'; (b) /t/ in 'toast; and (c) /k/ in 'come'

Representing the voiceless stop production by the Javanese students, examples showing the spectrograms of /p/ production at the medial position and /k/ production at the initial position are provided in Figure 1. In Figure 1(a), we can see how /p/ is produced at the minimum VOT of 4 ms; Figure 1(b) shows how /t/ is produced at the average /t/ VOT of 47 ms; and Figure 1(c) is how /k/ is produced at the maximum VOT of 171 ms.

Figure 1(a) displays that even when preceded by a voiced consonant [ɹ] at the medial position, a full closure is created as seen from the light grey area of the spectrogram. The VOT is measured from the release or burst (Bu) point, which is shown by a dark area at the lower part of the spectrogram, and it ends when the voicing (Vo) of the vowel [ɜ] begins as shown by the spread dark grey area to the upper part of the spectrogram. The highlighted part in the spectrogram shows that the VOT area is in a very brief moment, measured at 4 ms. It also shows that even though at the medial position, there is a full airflow obstruction with no voicing, meaning that the value is positive.

Figure 1(b) presents a different environment of a voiceless stop at the initial position. It starts with a total closure (Cl) shown by the white area of the spectrogram. Suddenly, there is a dark area at the upper part, which is the burst (Bu), or a sudden release involving great energy, where [t] is released, marking the starting point of the VOT measurement. The dark area then fades into light grey, which marks the aspiration of [t], thus the symbol [h] comes along with [t]. As the voicing of the vowel [o] starts, the spectrogram becomes darker again, marking the end of the VOT area. Both the burst and the aspiration are considered as the VOT area, measured at 47 ms. Therefore, it is considered long-lagged.

Figure 1(c) shows another voiceless stop at the initial position. A total closure (Cl) is marked by the white area in the spectrogram. The VOT measurement starts at the release of [k] that creates a burst (Bu), marked by the sudden dark area in the spectrogram. The dark area then fades to a constant grey marking the aspiration of [k], hence the [^{h}] symbol. The aspiration ends when the voicing of the vowel [ɑ] begins, which also marks the end of the VOT area. Both the burst and the aspiration areas make the VOT area, as shown by the highlighted part, and measured at 171 ms, with obviously a positive value and considered as long-lagged.

VOTs of Javanese Students' English Word-Initial Voiceless Stops

With 319 data, word-initial voiceless stops show the same tendency as the overall voiceless stop data. The VOT means of each stop for both male and female students are only slightly longer by 2–8 ms compared to the overall voiceless stops, with /t/ by female students having the smallest gap and /k/ by female students having the biggest gap. The VOT pattern of both male and female students remains the same, which is /k/ > /t/ > /p/.

Table 2. VOTs of Javanese Students' English Word-Initial Voiceless Stops

Sound	F			M		
	n	$\bar{x}$	Range	n	$\bar{x}$	Range
p_in	54	32	4–105	54	34	8–117
t_in	52	44	11–127	51	55	8–159
k_in	54	66	17–149	54	68	32–171

As we can see from Table 2, according to the VOT means, both male and female students' /p/ is short-lagged, and /t/ and /k/ are long-lagged. Male students still have longer VOT for each stop than female students, and /t/, again, is the one with the biggest gap among the three stops by 11 ms. This is in line with the previous study which found that /t/ is the most gender-related among the voiceless stops (Korkmaz & Boyaci 2018).

The ranges are also almost the same as the ones in the overall voiceless stop data, with over 100 ms range for each stop consonant. The male students' minimum and maximum data have higher VOT values than female students. The minimum and maximum values for female students' /p/ and /k/ and male students' /t/ are exactly the same as the ones in the overall voiceless stops. The maximum values for male students' /p/ and /t/ are the same as well. This could happen because 319 out of 514 data are word-initial voiceless stops.

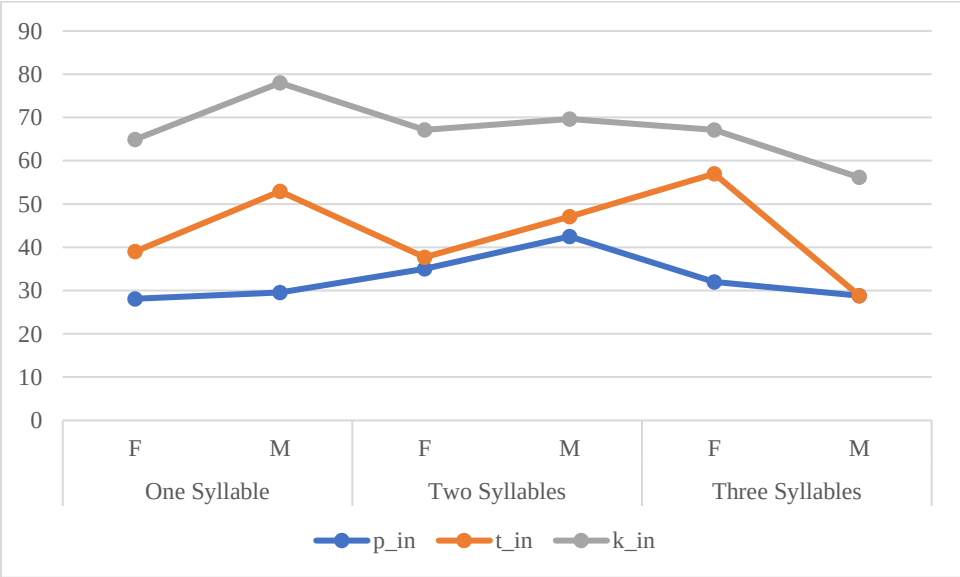


Figure 2. VOTs on Syllable-Basis of Javanese Students' English Word-Initial Voiceless Stops

If we take a look at the syllable basis, the VOTs are quite fluctuating. Figure 2 reveals that the notion that 'voiceless stops in one-syllable words should have longer VOT than those in more than one-syllable words' only applies to female students' /t/ and male students' /t/ and /k/. This finding indicates that the syllable number might not be the only factor affecting the Javanese students' VOTs of voiceless stops at the initial position. Since the word list used as the research instrument involves different environments with combinations of stressing positions and vowels, these factors are found to also affect VOT.

In the one-syllable word 'pet', /p/ is followed by a short vowel and a consonant; in the two-syllable word 'parrot', /p/ is followed by a short vowel, and the word is stressed in the first syllable; and in the three-syllable word 'personal', /p/ is followed by a long vowel and a

consonant, and the word is stressed in the first syllable. All the words are stressed in the first syllable and have short vowels but only the three-syllable word has a long vowel. According to the /p/ line, the VOT patterns are two syllables > one syllable > three syllables for female students and two syllables > three syllables > one syllable for male students, with insignificant gaps between one- and three-syllable words for both male and female students. From these patterns, we see that 'parrot' has the longest VOT for /p/, and it is hard to conclude the reasons behind the different patterns other than different environments.

In the one-syllable word 'toast', /t/ is followed by a long vowel and a double coda; in the two-syllable word 'toothbrush', /t/ is followed by a long vowel and a consonant, and the word is stressed in the first syllable; and in the three-syllable word 'tournament', /t/ is followed by a short vowel and a consonant, and the word is stressed in the first syllable. The three are all stressed in the first syllable and have long vowels but only the three-syllable word has a short vowel. According to the /t/ line, the VOT patterns are both one syllable > two syllables > three syllables for male and female students. The similar environment between one and another makes the pattern stable.

In the one-syllable word 'come', /k/ is followed by a short vowel and a consonant; in the two-syllable word 'comma', /k/ is followed by a long vowel, and the word is stressed in the first syllable; and in the three-syllable word 'composer', /k/ is followed by a short vowel and a consonant, the word is stressed in the second syllable. The three words have short vowels except for the two-syllable word that has a long vowel and are stressed in the first syllable except for the three-syllable word that is stressed in the second syllable. According to the /k/ line, the VOT patterns are one syllable > two syllables = three syllables for female students and one syllable > two syllables > three syllables for male students. From these patterns, we know that 'come' has the longest VOT for /k/. The environment between the one- and two-syllable words for /p/ and /k/ are actually similar where both 'pet' and 'come' have a short vowel and they are followed by a consonant. Yet, the result shows the opposite.

Based on the elaboration above, it seems that the notion of 'a stop in one syllable word has a longer VOT than in more syllables' only applies to words whose environments are similar. Like what happened to /t/, the three words are all stressed in the first syllable, and all the vowels are followed by consonants. If the environments are different like what occurs to /p/ and /k/, the notion does not apply.

VOTs of Javanese Students' English Word-Medial Voiceless Stops

At the medial position, 185 data also show the same tendency as the one at the initial position. The average VOT of each stop consonant is shorter than the ones at the initial position. The differences from the VOT means at the initial position are quite significant. It means the study finding matches the theory that VOT at the initial position is longer than at the medial position (Carley & Mees 2020; Collins et al. 2019; Fromkin et al. 2014; O'Grady & Archibald 2016; Zsiga 2013). The differences between the VOT means of each stop for both male and female students from the previous category are quite big, from 6 to 20 ms, but on exactly the same stop categories, with /t/ by female students being the shortest and /k/ also by female students being the longest. Moreover, the VOT pattern is still /k/ > /t/ > /p/.

Table 3. VOTs of Javanese Students' English Word-Medial Voiceless Stops

Sound	F			M		
	n	$\bar{x}$	Range	n	$\bar{x}$	Range
p_me	36	23	4–58	36	27	7–72
t_me	28	38	7–149	13	48	16–51
k_me	36	46	19–85	36	51	23–92

From Table 2 we can see that, as also occurs at the initial position, both male and female students' /p/ is short-lagged, and /t/ and /k/ are long-lagged. Male students, again, have longer VOT for each stop than female students, and /t/ remains to have the biggest gap among the three stops by 10 ms. It means the previous study's finding that /t/ is the most gender-related among the voiceless stops (Korkmaz & Boyaci 2018) remains true at the medial position.

Despite all the similarities with the previous category, the ranges show a difference. Overall, the ranges at the medial position are shorter for /p/ and /k/ with quite big gaps, but not /t/. /p/ and /k/ have VOT gaps of 65 and 69 ms at the maximum, while /t/ still has over 100 ms, for both male and female students. Male students' /t/ range is slightly shorter than the previous category but female students' /t/ range is bigger in this category.

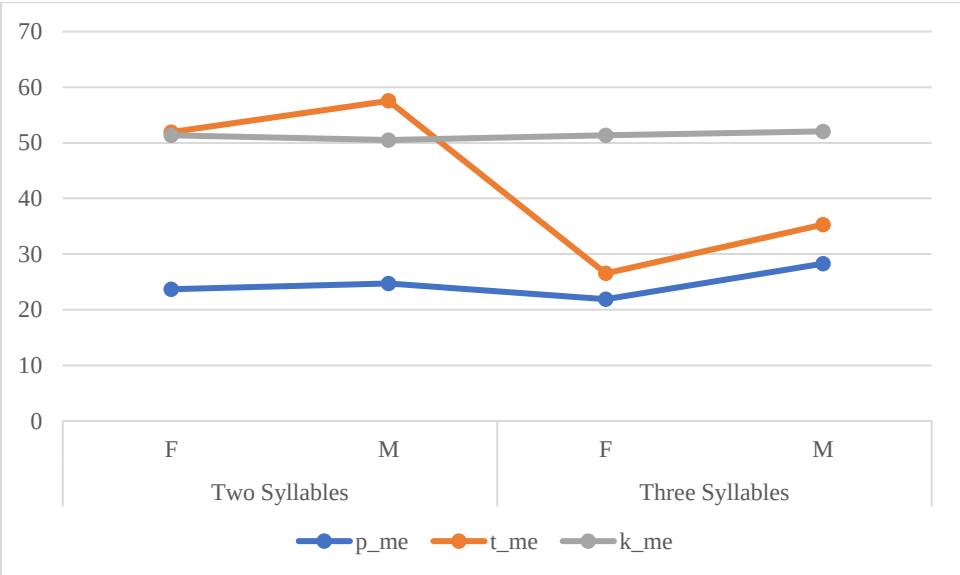


Figure 3. VOTs on a Syllable Basis of Javanese Students' English Word-Medial Voiceless Stops

On a syllable basis, the VOTs in the medial position category show different characteristics compared to the ones in the initial position category. In this category, the lines in Figure 3 tend to be flat for /p/ and /k/. With only little gaps, the VOT of female students' /p/ in two-syllable words is longer than in three-syllable words, and /k/ VOT shows no difference between the two-syllable numbers. Meanwhile, male students' VOTs for /p/ and /k/ are longer in two-syllable words than in three-syllable words. /t/ VOT is the only one showing no difference between male and female students. At the medial position, the environments are more complicated than words in the initial position category. Other than being influenced by the word stress and vowel type, they are also affected by the preceding sound.

In the two-syllable word 'super', /p/ is preceded by a long vowel and followed by a short vowel and a consonant; and in the three-syllable word 'carpenter', /p/ is preceded by an approximant

/ɹ/ and followed by a short vowel and a consonant. The two words are stressed in the first syllable but they have different preceding sounds, one vowel and one consonant. According to the /p/ line, the VOT patterns are two syllables > three syllables for female students and three syllables > two syllables for male students.

In the two-syllable word 'dirty', /t/ is preceded by an approximant /ɹ/ and followed by a short vowel and a consonant; and in the three-syllable word 'interview', /t/ is preceded by a nasal /n/ and followed by a short vowel and a consonant. Both the two words are stressed in the first syllable and they have different preceding sounds, where both are voiced alveolar consonants, but one is nasal and one is approximant. According to the /t/ line, the VOT patterns of male and female students are the same, which is two syllables > three syllables. The stable pattern might occur due to the similar environment of the consonant, although they are two different phonemes.

In the two-syllable word 'escape', /k/ is preceded by a fricative /s/ and followed by a long vowel and a consonant; and in the three-syllable word 'orchestra', /k/ is preceded by an approximant /ɹ/ and followed by a short vowel and a consonant. The two words are both stressed in the first syllable and preceded by alveolar consonants, one is voiceless fricative and one is voiced approximant. Moreover, 'escape' might also be pronounced with /s/ as part of the second syllable, making /k/ the second onset of the syllable. According to the /k/ line, the VOT patterns are two syllables = three syllables for female students and three syllables > two syllables for male students, with only an insignificant gap. The different patterns might be caused by the different environments of the preceding sounds.

Based on the explanation in the three previous paragraphs, syllable numbers seem to ambiguously affect the production of word-medial voiceless stops. What happened to /t/ is one thing but it might correlate with the disproportional data between male and female students. However, one thing to conclude is that other than the stressing position and the following vowel, the preceding sound affects the production of word-medial voiceless stops of the Javanese students.

Phoneme Shifts in Javanese Students' English Voiceless Stops

As previously mentioned, from the total of 540 data, only 514 data were measured using Praat. The 26 data were excluded due to phoneme shifts that only occur to /t/ both at initial and medial positions. Of the 26 shifts, 16 are produced by male students and 10 by female students, making it in line with the previous study where /t/ is the most gender-related among the three voiceless stops (Korkmaz & Boyaci 2018), and it applies not only to VOT but also to English /t/ variations.

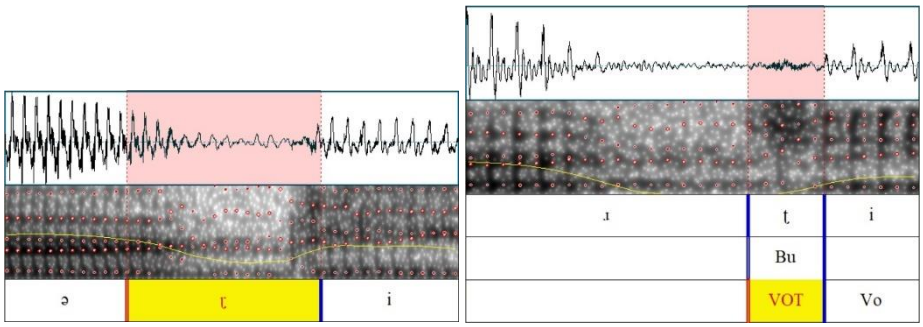
At the initial position, there are only 5 occurrences of phoneme shifts. Of the 5 data, 3 were done by male students and 2 by female students. Meanwhile, at the medial position, 21 phoneme shifts were found. Among the 21 occurrences, 13 were produced by male students and 8 by female students. The phoneme shifts found at the initial position are all due to pronunciation errors, either because of the slip of the tongue or a lack of English capability, the pronunciation errors also appear at the medial position but most of the word-medial phoneme shifts are English /t/ variations that are normally pronounced by native speakers.

Phoneme /t/ at both the initial and medial positions are found to be realized as [θ]. There is only 1 datum produced by a female student to the one-syllable word-initial 'toast' and another

produced by the same student to the two-syllable word-medial 'dirty'. /θ/ is a phoneme in English that is not part of the Javanese phoneme. That is why it substituting /t/ might occur due to a tongue slip because /t/ and /θ/ are both English dental consonants. This pronunciation error was also found in the production of English /t/ by Makassarese and Buginese students (Utami et al. 2017).

Another realization of /t/ at both the initial and medial positions is [tʃ]. There are 4 occurrences of /t/ realized as [tʃ] at the initial position, 1 by a female student and 3 by male students to the three-syllable word 'tournament'. At the medial position, it occurs to the three-syllable word 'interview.' /tʃ/ is a phoneme existing in both English and Javanese. The English /t/ substitution to [tʃ] might be caused by an unintentional slip of the tongue due to the perception that English /t/ is fricative produced with aspiration and when it is actually plosive since both are produced at the alveolar area.

A variation of /t/ which also only occurs at the medial position is the flap [ɾ] to the two-syllable word 'dirty'. There are a total of 4 occasions of this type of substitution, with 3 produced by female students and 1 by male students. The occurrences are possible because [ɾ] is produced by briefly and lightly slapping the tongue tip to the alveolar ridge, causing it to sound like alveolar /d/ (Gussenhoven & Jacobs 2017).



(a) (b)

Figure 4. /t/ realized as (a) [ɽ] and (b) as [t] in 'dirty'

Another variation of word-medial /t/ that occurs to 'dirty' is [ɽ]. There are 5 realizations of this type, with 4 produced by male students and only 1 produced by a female student. Since [ɽ] is basically the retroflexed version of [ɾ], the way they are produced is similar. Only, the flap in [ɽ] is thicker with the tongue curl. Figure 4(a) shows the production of [ɽ] by a student in a spectrogram. Since [ɽ] only involves a flap, there is no total closure caused by it. As a consequence, the voicing of the preceding syllable's vowel remains, causing a light grey area before the flap occurs, as marked by a sudden darker area before entering the next vowel. In the production of [ɾ], the light flap does not cause the spectrogram as dark as the thick flap of [ɽ].

The word 'dirty' originally involves the approximant /ɹ/ but native English speakers, prominently Americans and Australians, might pronounce the /t/ with [ɾ]. The retroflex flap [ɽ] possibly appears in the Javanese students' production of English /t/ due to the influence of Javanese retroflex /ʈ/. In fact, there is one occasion where a male student pronounces /t/ in 'dirty' with [t]. Unlike [ɽ], /ʈ/ causes a total airflow closure, hence causing a plosive sound that is thicker than the flap in [ɽ]. Therefore, as a comparison, it can be seen from the spectrogram in Figure 4(b) that the dark area of /ɹ/ at the beginning fades to be lighter and lighter. The

soundwave also seems to be flat, which is flatter than the one in Figure 4(a) when the flap of [ɽ] is about to happen. Back to Figure 4(b), as /ʈ/ appears, there is a sudden dark area that is darker than the flap of [ɽ] in Figure 4(a).

The last type of /t/ variation found in this study is [n], which occurs at the medial position to the three-syllable word 'interview'. Out of 10 occurrences, 8 were produced by male students and only 2 were produced by female students. In this case, /t/ totally disappears, and [n] which is actually the coda of the preceding syllable plays a double role as also the onset of the syllable substituting /t/. This variation is also normally done by American and Australian native English speakers when they have daily conversations.

Javanese Identity Shown through Phonological Transfers and Phoneme Shifts

Language use could reflect someone's identity. Phonological transfers and phoneme shifts from one's first language that appear in their second or third language might show how someone uses their languages and reveal how they perceive these languages. In the case of this research, the authors assume that Javanese phonological transfers and phoneme shifts in the production of English voiceless stops might show the Javanese students' language attitude amidst the multilingual environment in Indonesia.

The Javanese students already follow the VOT pattern k > t > p and there is not much difference between male and female students' English voiceless production. However, other than the early findings that women tend to use more standard language than men (Sunderland 2006), the difference might occur due to male students' average English ability which was slightly better than female students. Most students were at a B2 level when the screening test was carried out, with some female students at a B1 level and some male students at a C1 level; meanwhile, no male students were at a B1 level, and no female students were at C1 level.

The English screening test result might also explain why more male students imitate the English native speakers' accents than female students, causing phoneme shifts in both initial and medial positions. On the one hand, these male students are possibly more exposed to the standard British English-speaking style, hence the perception of the thick aspiration that comes with /t/ in the initial position that is then falsely produced as fricative [tʃ]. On the other hand, they are also probably influenced by the American and Australian English accents, causing phoneme shifts in the medial position: the disappearance of /t/ replaced by /n/, which is actually the coda of the previous syllable, and the substitution of the approximant /ɹ/ with the flap [ɾ].

Interestingly, Javanese phonological transfers occur during the process of replacing the approximant /ɹ/ with the flap [ɾ] in the medial position. Five students mistakenly produced the retroflex flap [ɽ] instead of [ɾ]. The existence of a student mispronouncing /t/ in 'dirty' with [ʈ] suggests that this shift, along with the retroflex flap [ɽ] replacing [ɾ], is influenced by the Javanese [ʈ].

Even though it occurs in 6 out of 180 English /t/ production in the medial position, further research is needed. This phenomenon would be better to be brought to a deeper level, for example, whether the students consciously or unconsciously produce it. If it is due to tongue slip, then the cause is probably because they are still English learners. If they willingly use their Javanese accent, or even when they do it unintentionally, regardless of their English ability, the language attitude they have towards this kind of phoneme shift is still worth to investigate.

Other plosives such as /b/, /d/, and /g/, and fricative /dʒ/ could also be examined since they carry Javanese phonological quality whose phonetic characteristics are still limitedly examined.

This research could only lead to the above assumptions so far. More linguistic evidence is needed to see the dynamics of language usage among the younger generation in Indonesia. Based on the 2020 long-form census, the Statistics Indonesia's latest release concerning the language usage in Indonesia (Purbowati et al. 2024) discloses that although at least 42% of the total population in Indonesia still speaks Javanese, vernacular language usage decreases by almost 5%, and the younger the generation, the less vernacular language is used by Indonesian people. Since these people live in a multilingual environment, first-language phonological transfers, including those that occur in English, could be an indicator of their attitude toward their language usage.

CONCLUSION

This study found that regardless of position and gender, based on the overall VOT means, the pattern remains the same, which is /k/ > /t/ > /p/. In other words, the more posterior the pronunciation location, the longer the VOT. Regardless of the position, male students have longer VOT than female students. Voiceless stops at the initial position have longer VOT than at the medial position. There is an indication that the syllable number may affect the pattern at the initial position, and it is less influential at the medial position because the stop environments, such as the stressing position, the vowel type, and the preceding sound of word-medial stops, have more influence on the VOT. To conclude, male students have more variations of /t/ and seem to be more Javanese-influenced with more occurrences of [ʈ] than female students although they produce longer VOT.

Future research is suggested to consider other possible transfers that might occur due to distinct phonological characteristics such as in Javanese /b/, /d/, /g/, and /dʒ/. On-site recording using more appropriate tools and carrying out more advanced statistical analysis will be necessary to produce quality data and more convincing results. With this adequate setting, more robust research results are expected to bring a more reliable interpretation of the language dynamics in Indonesia, which can be used as a basis for language learning policy in Indonesia.

ACKNOWLEDGEMENT

N/A

REFERENCES

Ambalegin, & Hulu, F. 2019. "EFL Learners' Phonological Interference of English Articulation." *BASIS* 6 (2), 145–155. DOI: https://doi.org/10.33884/basisupb.v6i2.1415.

Bolton, K. 2006. "World Englishes Today." In *The Handbook of World Englishes*, edited by B. B. Kachru, Y. Kachru, & C. L. Nelson. Blackwell Publishing.

Buchstaller, I. & G. Khattab. 2013. "Population Samples." In *Research Methods in Linguistics*, edited by R. J. Podesva & D. Sharma. Cambridge: Cambridge University Press.

Carley, P., & I. M. Mees. 2020. *American English Phonetics and Pronunciation Practice*. Routledge.

Chodroff, E., A. Golden, & C. Wilson. 2019. "Covariation of Stop Voice Onset Time Across Languages: Evidence for a Universal Constraint on Phonetic Realization." *Acoustical Society of America* 145 (1). DOI: https://doi.org/10.1121/1.5088035.

Collins, B., I. M. Mees, & P. Carley. 2019. *Practical English Phonetics and Phonology* (Fourth Edition). Routledge.

Fromkin, V., R. Rodman, & N. Hyams. 2014. *An Introduction to Language* (Tenth Edition). Cengage Learning.

Genetti, C. 2014. *How Languages Work*. Cambridge: Cambridge University Press.

Gussenhoven, C. & H. Jacobs. 2017. *Understanding Phonology* (Fourth Edition). Routledge.

Kachru, B. B., Y. Kachru, & C. L. Nelson (eds.). 2006. *The Handbook of World Englishes*. Blackwell Publishing.

Kirkpatrick, A. 2010. *The Routledge Handbook of World Englishes*. Routledge.

Korkmaz, Y. & A. Boyaci. 2018. "Analysis of Speaker's Gender Effects in Voice Onset Time of Turkish Stop Consonants." *6th International Symposium on Digital Forensic and Security (ISDFS)*, 1–5. DOI: https://doi.org/10.1109/ISDFS.2018.8355341.

Ladefoged, P. & S. F. Disner. 2012. *Vowels and Consonants* (Third Edition). Wiley-Blackwell.

Marsono. 2018. *Fonetik*. Yogyakarta: Gadjah Mada University Press.

———. 2019. *Fonologi*. Yogyakarta: Gadjah Mada University Press.

Mesthrie, R. 2004. "Synopsis: Phonological Variation in Africa." In *A Handbook of Varieties of English 1: Phonology* (Vol. 1), edited by E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, & C. Upton. Berlin: Mouton de Gruyter.

Nelson, C. L., Z. G. Proshina, & D. R. Davis (eds.). 2020. *The Handbook of World Englishes* (Second Edition). John Wiley & Sons, Inc.

O'Grady, W. & J. Archibald. 2016. *Contemporary Linguistic Analysis: An Introduction* (Eight Edition). London: Pearson.

Oh, E. 2019. "Effects of Gender on the Use of Voice Onset Time and Fundamental Frequency Cues in Perception and Production of English Stops." *Linguistic Research* 36 (1), 67–89. DOI: https://doi.org/10.17250/khisli.36.1.201903.003.

Purbowati, A., D. Handiyatmo, D. Trisnani, E. Zoraya, I. Fahmi, N. A. Ramadhani, P. R. A. Padini, S. Wahyuni, & U. R. Sitorus. 2024. *Profil Suku dan Keragaman Bahasa Daerah Hasil Long Form Sensus Penduduk 2020*. Badan Pusat Statistik. https://www.bps.go.id/id/publication/2024/12/12/6feb932e24186429686fb57b/profil-suku-dan-keragaman-bahasa-daerah-hasil-long-form-sensus-penduduk-2020.html.

Senowarsito & S. N. Ardini. 2019. "Phonological Fossilisation of EFL Learners: The Interference of Phonological and Orthographic System of L1 Javanese." *3L: The Southeast Asian Journal of English Language Studies* 25 (2), 74–85. DOI: https://doi.org/10.17576/3L-2019-2502-06.

Soetjahyati, O. S. 2018. "The Acoustic Characteristics of Stop Sounds Produced by East Javanese Students of English Studying at University of 17 Agustus 1945 Surabaya." Surabaya: Universitas 17 Agustus 1945. http://repository.untag-sby.ac.id/id/eprint/892.

Statistics Indonesia. 2015. "Mengulik Data Suku di Indonesia." Statistics Indonesia. https://www.bps.go.id/news/2015/11/18/127/mengulik-data-suku-di-indonesia.html.

Suku Bangsa. 2017. "Suku Bangsa." Indonesia.go.id. https://indonesia.go.id/profil/suku-bangsa/kebudayaan/suku-bangsa, December 3.

Sunderland, J. 2006. *Language and Gender: An Advance Resource Book*. Routledge.

Thurgood, E. 2004. "Phonation Types in Javanese." *Oceanic Linguistics* 43 (2), 277–295. https://www.jstor.org/stable/3623359.

Utami, D. H., B. Wello, & H. Atmowrdoyo, H. 2017. "The Phonological Interference of Students' First Language in Pronouncing English Sounds (A Case Study on Buginese and Makasarese Students)." *ELT Worldwide* 4 (2), 205–212. DOI: https://doi.org/10.26858/eltww.v4i2.4414.

Wedhawati, W. E.S. Nurlita, S. Nardiati, Herawati, R. Sukesti, Marsono, E. Setiyanto, D. Sabariyanto, S. Arifin, Sumadi, & Laginem. 2001. *Tata Bahasa Jawa Mutakhir*. Jakarta: Pusat Bahasa Departemen Pendidikan Nasional.

Yu, V. Y., L. F. D. Nil, & E. W. Pang. 2014. "Effects of Age, Sex, and Syllable Number on Voice Onset Time: Evidence from Children's Voiceless Aspirated Stops." *Language and Speech* 58 (2), 152–167. DOI: https://doi.org/10.1177/0023830914522994.

Zen, E. L. 2018. "On the Acquisition of English Voiceless Stop VOT by Indonesian-English Bilinguals: Evidence of Input Frequency." *K@ta* 20 (2), 45–52. DOI: https://doi.org/10.9744/kata.20.2.45-52.

Zsiga, E. C. 2013. *The Sounds of Language: An Introduction to Phonetics and Phonology*. Wiley-Blackwell.