

Comparative Study of Voice Onset Time in English Word-Initial Stop Consonants Produced by Uzbek and American Speakers of English

KHIKMATULLO URAZBAEV
Westminster International University in Tashkent,
Uzbekistan
kurazbaev@wiut.uz

SHERZOD SUKHROBBEKOV
Chonnam National University,
South Korea

ABSTRACT

The purpose of this study is to shed light on a comparative study of Voice Onset Time (VOT) values produced by Uzbek speakers of English and native speakers of American English. The study aimed to determine significant differences in terms of the VOT value in pronunciation of voiced and voiceless stop consonants in word-initial positions. Scholarly sources on the comparative study of Uzbek and English phonetics suggest that most of the Uzbek stop consonants possess similar qualities to those of English stop sounds (Abduazizov 1992; Yusupov 2013). A descriptive comparison of Uzbek and English consonants demonstrates the difference between the two languages in isolated forms rather than in speech. This study is motivated by the fact that little is known about recent studies on the instrumental comparative study of English plosives produced by Uzbek speakers of American English. In addition, so far, no studies have been carried out in VOT analysis of English sounds produced by Uzbek speakers. The research was conducted through voice recording and analyses of data collected from Uzbek speakers and American native speakers. Recordings of 3240 overall tokens were collected and analyzed using Praat phonetic software. The mean VOT values obtained from both subject groups have been compared and statistically tested using the Independent Samples t-test. The research findings show that Uzbek subjects' and American native subjects' VOT values differ significantly. However, there are a few words (9 out of 52) where the differences in the VOT values of the two subject groups are not statistically significant. The research presents implications for teaching pronunciation of English plosive sounds in word-initial positions where a stop consonant followed by a vowel sound to Uzbek speakers of English by taking into account the periodic measurement of consonants. In addition, there are implications for future research that might explore other features of VOT produced by Uzbek language speakers.

Keywords: Voice Onset Time; stop consonants; voiced; voiceless consonants; Uzbek

INTRODUCTION

In English phonetics, consonants have been the object of studies for many decades and studied from many angles. One of them in which English consonants and consonants of other languages have been studied is the Voice Onset Time (VOT), introduced by Lisker and Abramson (1964). Afterwards, VOT has become one of the necessary means of measuring the voicing of stop consonants in English and other world languages, especially in word-initial positions.

English and Uzbek originate from very different language families. Uzbek belongs to the Altaic family of languages, whereas English falls under the umbrella of the Indo-European family of languages, which assumes contrastive rather than comparative analysis of phonological characteristics of both languages. In Uzbekistan, very few scholars addressed the contrastive phonological analysis of Uzbek and English sounds (Abduazizov, 1981,1992; Yusupov, 2013). Also, a review of the literature published in Uzbekistan on comparative and contrastive phonology of English and Uzbek revealed that there are no studies conducted in the comparative or contrastive analysis of VOT values of English stop consonants produced by Uzbek speakers of English.

Consonants and the Voice Onset Time are considered to be two of the cues that determine L1 accent and L2 learners' native speaker accent production (Munro & Derwing, 1995; Schoonmaker-Gates, 2015). Due to reasons similar to the context of Arab countries, as mentioned by Mustafa (2019), pronunciation has been underestimated in EFL classrooms in the context of Uzbekistan. In contrast to Arab EFL contexts, this is caused by a lack of instrumental research studies on actual Uzbek English learners. Shedding light on Voice Onset Time values produced by Uzbek English speakers may enhance English pronunciation instruction and provide an impetus for research-informed pronunciation teaching in the Uzbek context. Thus, the aims of the given research are as follows:

1. To provide a comparison of the VOT values of the English voiceless and voiced consonants in word-initial positions produced by native American English speakers and Uzbek speakers of English;
2. To identify the differences in VOT patterns in pronunciation of English voiceless and voiced stop consonants by Uzbek speakers of English.

LITERATURE REVIEW

Voice Onset Time or VOT is one of the critical features prevalent in the production of consonants. It is the “duration of the time between the release of a plosive and the beginning of vocal folds’ vibration” (Bilá & Eddy, 2012, p. 49). There are three types of VOT: zero VOT (plosive release and vibration happen together), positive VOT (the plosive is released and vocal folds start vibrating with a delay), and negative VOT (vocal folds start vibrating and then the plosive is released). Logically, a voiceless/fortis plosive has positive VOT, while a voiced/lenis plosive has zero VOT.

VOT is an immediate parameter, which can be distinguished in terms of time between the discharge of the full oral constraint for producing the voice in a plosive manner and the onset vibrations of the glottal (Lisker & Abramson, 1964). In many languages, there are stops [b], [d], [g], which are voiced, and [p], [t], [k], which are voiceless. Moreover, one of the key indicators of voiceless and voiced plosive consonants are their VOT values (Olson, 2017).

In producing plosive sounds, the upper teeth contact the tongue, and these stops are referred to as dental stops. Bilabial stops are the consonants produced by the oral constriction provided by the upper and lower lips. Another type of consonant is velar stops, formed by the contact of the back of the tongue with the soft palate (McMahon, 2002). In the Uzbek language as in the English language, [p] and [b] are bilabial consonants, [k] and [t] are dental, and [k], and [g] are velar consonants.

As mentioned above, Lisker and Abramson (1964) were the pioneers in investigating the VOT of NSs of American English. After their contributions, many studies have been conducted regarding VOT of native speakers (NSs), bilingual speakers, and non-native speakers (NNSs) of various languages (Taehong & Ladefoged, 1999; Fowler et al., 2006; Olmstead et al., 2013; Kiliç, 2018). Further studies suggest that VOT plays a vital role in distinguishing individual talker differences (Allen et al., 2003; Allen & Miller, 2004).

Lisker and Abramson (1964) in their research established that the VOT values of the stop consonants [t], [d], [p], [b], [k], and [g] were 70, 5, 58, 1, 80, and 21 milliseconds (ms) respectively. In the study conducted by Klatt (1975), it was found that VOT values of the stop consonants [p], [t], and [k] were 47, 65, and 70 ms respectively, whereas the VOT values of the stop consonants [b], [d], and [g] were found to be 11, 17, and 27 ms respectively.

Caruso and Burton (1987) studied the VOT values of 8 native English speakers and established that the VOT values of the stop consonants [b], [p], [d], [t], [g], and [k] were as follows: 19.7, 62.5, 21.4, 71.9, 21.4, 74.8 ms respectively.

Findings regarding the VOT values of stop consonants produced by native English speakers are summarised in Table 1 below.

TABLE 1. Vot values of english stop consonants reported

Authors	Stop Consonants and Reported VOT Values						Subjects
	[p]	[b]	[t]	[d]	[k]	[g]	
Lisker & Abramson (1964)	58	1	70	5	80	21	AmE Speakers
Klatt (1975)	47	11	65	17	70	27	AmE Speakers
Caruso & Burton (1987)	62.5	19.7	71.9	21.4	74.8	35.2	Native English speakers
Kessinger & Blumstein (1997)	85	15	100	20	n/a	n/a	Native English speakers
Macleod & Stoel-Gammon (2005)	87.9	19.8	87.9	19.8	n/a	n/a	Canadian English Speakers

The studies given in Table 1 are not comparative studies and do not provide a comparative analysis of the mean VOT values produced by NSs and NNSs. Nevertheless, the mean VOT values found in the above studies will be used as a baseline to compare the findings in this research.

The research findings conducted by Kessinger & Blumstein (1997) suggested that an increase in the speech rate shortens the VOT value. Klatt (1975) and Higgins et al. (1998) found that the sound which follows the stop consonants affects the VOT value of the word-initial stop consonant. According to Klatt (1975), the VOT value increases in length when sonorant consonants follow the stop consonant. Moreover, Klatt (1975) also reported that when followed by high vowels, the length of VOT increases, whereas the VOT of the stop consonants is shorter when low, open vowels follow them.

Uzbek is a Turkic language, the official language of the Republic of Uzbekistan, and a regional language in eight Northern provinces (viloyats) of Afghanistan. Also, it is wide-spread in Tajikistan, Kyrgyzstan, Kazakhstan, Turkmenistan, Russia, Turkey and other countries (Shoabdurahmonov et al., 1980). Uzbek is a dialectal language, which allows it to be included in different subgroups (Nabieva, 2005). It is the mother tongue and the primary language for most Uzbeks globally (Tursunov et al., 1995, p. 138).

In the modern Uzbek language, there are 24 consonants. As in other units of layer, the main features of consonants can be sub-divided into correlative and non-correlative features (Mirtozhiev, 2013, p. 97). These features, which are existent in one of the pair consonants, are called the correlative feature. In the Uzbek language there are 8 pairs of voiced and voiceless consonants: [b]-[p], [d]-[t], [k]-[g], [s]-[z] [j]-[ch], [j]-[sh], [gʻ]-[q], [v]-[f] (Nurmonov, 1990; Nurmonova, 1998; Mirtozhiev, 2013).

In comparing and contrasting Uzbek consonants with their counterparts in the English language, they are pretty often distinguished according to their similarity, partial similarity, and difference (phonetic gap or lacuna).

The Uzbek language has the same number of voiced and voiceless stop consonants as the English language (Mirtozhiev, 2013). They are [p], [b], [t], [d], [k], and [g]. The literature review revealed no extensive quantitative research on the Voice Onset Time of Uzbek consonants in word-initial positions. Mirtozhiev (2013, p. 97) states that stop consonant [p] in

the Uzbek language may occur differently. In word-initial positions, the [p] sound in the length of production is similar to [b] and pronounced in 40-60 ms.

Uzbek sound [b] is one of the most active sounds in word-initial positions and is produced from 150 ms to 200 ms depending on the preceding sounds. On the other hand, Mirtozhiev (2013, p. 97) notes that [b] is relatively inactive in word-medial positions, and it takes 50 and less ms to produce this sound.

In the Uzbek language, stop consonant [d] occurs both in word-initial and word-final positions. In the word-initial positions, it actively occurs in the Uzbek language. In the word-initial positions, if the preceding sound is a short vowel, length can vary from 50 to 70 ms, whereas in the word-initial position before a long vowel sound, length varies from 110 to 150 ms (Mirtozhiev, 2013, p. 98).

Uzbek stop consonant [t] can occur virtually in all positions. The most distinguishing feature of the stop consonant in the Uzbek language is that it has a similar length, 60 ms in all word-initial, word-medial and word-final positions. In some cases, if the preceding sound is a vowel and [t] moves towards becoming the [d] sound, its pronunciation length may fall to 50 ms (Mirtozhiev, 2013, p. 98).

In the Uzbek language, plosive stop consonant [g] is somewhat limited according to its position in words. In native Uzbek words, [g] is not observed in word-initial positions. It occurs mainly in borrowed words in the Uzbek language. According to Mirtozhiev (2013, p. 99), the length of the sound [g] in word-initial positions varies between 150 and 170 ms. In word-medial positions, the length of the [g] sound may be at 40-60 ms.

Uzbek voiceless stop consonant [k] is not limited according to its position in words. In word-initial positions, its pronunciation is short. For instance, in words like [kam] (few), [kel] (come), [kosa] (saucer) and [kelin] (bride), the length of its pronunciation may vary from 35 to 40 ms (Mirtozhiev, 2013, p. 99). If the preceding vowel sound is short, its pronunciation length can be up to 70 ms. In word-medial positions, the length may vary from 60 ms to 75 ms.

Since most English language teachers favour American English pronunciation in Uzbekistan (Bakieva & Rustamova, 2020), the study aimed at conducting an instrumental comparative analysis of English stops produced by Uzbek speakers of English and stops produced by American English speakers.

METHODOLOGY OF THE RESEARCH

PARTICIPANTS OF THE STUDY

Convenience sampling was used in deciding the subjects of the study. Speakers of the Uzbek language for the study were chosen among the students of Gulistan State University (Uzbekistan). The total number of subjects was 10 (n=10) for Uzbek speakers of English. Since most of the subjects were students of a bachelor's degree course majoring in English language and literature, the age range of Uzbek speakers was 19 to 22.

Subjects who speak English as their mother tongue were the selected NSs of American English and who reside in South Korea and teach at universities and private schools in Chonnam. The age range of the participants is 30-50. In total, 10 (n=10) native English speakers took part in the voice recording session.

EXPERIMENTAL WORDS

Overall, each subject in each group (Group 1: American English (AmE) Speakers, n=10 and Group 2: Uzbek Speakers of English, n=10) was asked to pronounce 54 words during the voice recording process. However, for statistical significance purposes, it was decided to have 20

subjects (n=20) in the study as the times each experimental word was pronounced would accumulate tokens which would be enough for statistical analysis. Therefore, the study subjects were asked to pronounce each word three times with a slight pause in between them. Thus, in total, 3240 tokens produced by 20 subjects were analyzed in the research.

VOT value in the production of voiced stop consonants [b], [g], [d] and voiceless stops [p], [k], [t] was the object of the study. In each set of consonants, three groups were determined according to the position of the consonants and the preceding vowel sounds [i], [i:], [ɑ:], [ɑ], [u], and [u:]. Thus, all words' syllabic structure was consonant-vowel-consonant (C-V-C) (see Appendix for experimental words).

INSTRUMENTS

Several voice recording and analyzing products were employed in the acoustic recording and processing stages of the research.

The device used for an acoustic recording of the participants' voices in the context of Uzbekistan was Digital Voice Recorder CL-R30 650Hr. The samples were recorded using the 44.1 kHz setting on the voice recorder, and the output file was in 'wav' format, which is compatible with Praat software.

Samsung S9+ was used with the NSs of American English to record their voices. The recording setting was set to 44.1 kHz. Samples were initially recorded in '3gp' format and converted to 'wav' format to make them compatible with the Praat software.

Processing the recorded voices and measuring the VOT value was conducted using Praat Software version 6.1.08 (Boersma & Weenink 2019).

The numerical data obtained were further processed using MS Excel 2013 and IBM SPSS Version 17.0. MS Excel 2013 was used to calculate the VOT of each subject and to compute the mean VOT for each experimental word.

RESULTS AND DISCUSSION

In this section, the subjects of American English will be labelled as NSs (NS), and Uzbek speaker subjects will be labelled as NNSs (NNS).

VOT VALUE OF WORD-INITIAL STOPS IN HIGH VOWEL [i:]

VOT OF [p] AND [b]

In the experimental word "peat", the mean VOT value of NNSs was 27 ms. For most of the NNSs, the average VOT value ranged from 17 to 24 ms. In NS, the mean VOT value for the word "peat" was 60 ms. The standard deviation (SD) for NNSs was 13.2 and 16.1 for NSs. When the mean VOT value was statistically tested in SPSS for significance, it was found that there is a statistically significant difference in the mean VOT value in the word "peat".

The findings of the given research regarding American NSs are in line with the findings of Lisker and Abramson (1964) and Caruso and Burton (1987), all of whom established that the VOT values of the sound [p] in American English speakers are 58 ms and 62.5 ms respectively.

In the pronunciation of the word "peace", the mean VOT value for NNSs was 28.8 ms (SD=17), whereas, for NSs, the mean VOT value was 64 ms (SD=21).

An Independent samples t-test was conducted to compare the mean VOT value for NNSs and NSs. There was a significant difference in mean VOTs for NNSs (M = 29.02, SD = 17) and NSs (M = 64.17, SD = 21; $t(18) = -4.03, p = .000$, two-tailed).

The mean VOT value for the word “peek” for NNSs was 29.3 ms (SD=19.3) and for NSs was 77.5 ms (SD=23). Only one subject among the NNSs produced “peek” with the VOT value, close to NSs 72.9 ms.

The results of an independent samples t-test indicate that there is a significant difference ($p=.000$) in the mean VOT value for the sound [p] in the word “peek”.

In the pronunciation of the word “beet”, where the word-medial sound is long /i:/, the mean VOT value for NNSs was 101.6 ms (SD=49.2), whereas, for NSs, it was 43.4 ms (SD=40.3). Thus, the NSs of English findings do not match previous studies' findings (Lisker & Abramson, 1964; Klatt, 1975; Caruso & Burton, 1987; Kessinger & Blumstein, 1997; Macleod & Stoel-Gammon, 2005).

In the [b] sound production, Uzbek speakers tend to face mother tongue pronunciation interference, where they tend to produce the English stop consonant [b] in an Uzbek manner. As was mentioned in previous chapters, the VOT value of the Uzbek sound [b] varies from 150 to 200ms (Mirtozhiev, 2013).

An independent samples t-test showed that NNSs' pronunciation of the sound [b] significantly differs from NSs ($p=.010$).

For the word “beep”, NNSs pronunciation of the sound [b] has shown a positive mean VOT of 114 ms (SD=49.8). NSs' VOT value for the given sound was 19.6 ms (SD=17). It can be seen from both the data provided and the independent samples t-test carried out with tokens, there is a significant difference in Uzbek speakers' pronunciation of the sound [b] in terms of VOT value.

Also, in the experimental word ‘beak’, there was a significant difference between NSs and NNSs.

The NNSs' VOT value in all cases was significantly longer compared to those of NSs. For example, the mean VOT value for the sound [b] in the word “beak” was 100.4 ms (SD=41.5) for NNSs and 34.1 ms (SD=33) for NSs. The findings are in line with the previous studies of Klatt (1975), Caruso and Burton (1987) and Macleod and Stoel-Gammon (2005) if each native speaker is examined individually.

An independent samples t-test was conducted to compare the mean VOT value for NNSs and NSs. There was a significant difference in the mean VOTs for NNSs ($M = 100$, $SD = 41$) and NSs ($M = 34$, $SD = 33$; $t(18) = 3.95$; $p = .001$, two-tailed).

In the given experimental word, it was revealed that Uzbek speakers of English tend to produce a longer VOT value compared to NSs. This can be explained by mother tongue interference in the production of the experimental words.

VOT OF [t] AND [d]

In the word “teak”, where the stop consonant is followed by high vowel [i:], the mean VOT value for NNSs' pronunciation was 52.9 ms (SD=20), and for NSs, it was 85.7 ms (SD=17). A similar NSs VOT was reported by Macleod and Stoel-Gammon (2005). As for the Uzbek speakers, in Mirtozhiev's (2013) study, it was reported that the VOT of the sound [t] is equal to 50 ms, which is found in the given research in the case of English stop [t]. Thus, even though the mean VOT value in NSs was longer, there were subjects among NNSs, who produced the sound [t] with almost the same VOT value as NSs.

An independent samples t-test was conducted to compare the mean VOT value for NNSs and NSs. There was a significant difference in the mean VOTs for NNSs ($M = 52.9$, $SD = 20$) and NSs ($M = 85.7$, $SD = 17$; $t(18) = -3.89$; $p = .001$, two-tailed). According to the t-test results, there is a significant difference in the mean VOT value for the sound [t] in “teak”.

The mean VOT values for experimental words “tease” and “teat” showed 45.9 ms (SD=12.2) and 49.9ms (SD=18.1) for NNSs and 89.9ms (SD=22.6), 75.1 ms (SD=21) for NSs, respectively.

Individual examination of the mean VOT values shows that the research findings follow Lisker and Abramson (1964), Klatt (1975), Caruso and Burton (1987), Kessinger and Blumstein (1997), and Macleod and Stoel-Gammon (2005).

T-test results show a significant difference in the production of the word-initial sound [t] in both “tease” and “teat” by NNSs and NSs ($p = .001$, and $p = .010$, two-tailed).

The mean VOT values for the[d] sound in three different word tokens “dean”, “deed”, and “deep” with vowel sound [i:] were found to be 108.7 ms (SD=46), 86.5 ms (SD=34), and 104.3 ms (SD=47) for NNSs, respectively. The mean VOT values in the same word tokens for the NSs were found to be 51.7 ms (SD=42), 39.3 ms (SD=31.4), and 36.4 ms (SD=32.3), respectively.

A t-test was conducted for the mean VOT value in the given experimental words. The t-tests showed that there was a significant difference in the pronunciation of these words (“deep”, $p=0.001$, “deed” $p=0.005$ and “dean” $p=0.010$).

VOT OF [k] AND [g]

For the experimental words “keen”, “keep”, and “key”, mean VOT values for NNSs were found to be 82.9 ms (SD=18.9), 81.2 ms (SD=12.3) and 86.2 ms (SD=25.9), respectively. On the other hand, NSs produced the exact words with the mean VOT values of 111.9 ms (SD=38), 99.7 ms (SD=26.4) and 111.2 ms (SD=30.8), respectively.

An independent-samples t-test was conducted to compare the mean VOT value for NNSs and NSs. The t-test showed no significant difference in the mean VOT value for NNSs ($M = 81.2$, $SD = 12$) and NSs ($M = 99.7$, $SD = 26$; $t(18) = -2.01$; $p = .060$, two-tailed).

The research findings suggest that in producing the word “keep”, NNSs produce the same mean VOT value in the word-initial [k] sound as NSs do. Thus, the findings suggest that Uzbek speakers can pronounce the [k] sound without mother tongue interference.

In the production of the [k] sound in word-initial position, in the word “keen”, NNSs produce almost the same mean VOT value as the NSs. However, NSs’ mean VOT values are slightly longer compared to NNSs.

Next, in the production of the word “key”, the mean VOT value of the word-initial stop consonant [k] was found to be 86 ms (SD=25) in NNS and 111 ms (SD=30) in NSs.

The highest mean VOT value for the sound [k] found in NNS was equal to 143 ms and 185ms in the NS group. Thus, the overall trend shows that the mean VOT value in native speaker pronunciation is longer than the mean VOT value in Uzbek speaker production.

Nevertheless, a t-test conducted with the given word and the mean VOT value for NNSs and NSs has shown no statistically significant difference ($t(18) = -1.96$; $p = .065$, two-tailed).

For [k] in word-initial positions where the proceeding sound is high vowel sound /i:/, it was found that Uzbek speakers’ production of the [k] sound is similar to NSs.

According to the findings of Lisker and Abramson (1964), Klatt (1975), and Caruso and Burton (1987), the VOT value of the sound [k] ranges from 70 ms to 80 ms. However, the mean VOT values were higher in the given research than those reported in previous studies. Only close individual examination of each NS can show that the mean VOT values ranged from 70 ms to 100 ms. Nevertheless, the research revealed that the VOT value of NNSs was close to the findings of Lisker and Abramson (1964), Klatt (1975) and Caruso and Burton (1987), which supports the statistical results of this study. NNSs production of VOT value within the range of 70 to 90 ms can be explained by the fact that English stop sound [k] is pronounced similarly to Uzbek sound [k] in different contexts. Moreover, in the example of

the [k] sound, positive mother tongue interference can be observed. In the research conducted by Mirtozhiev (2013), it was found that the VOT value of the [k] was 70 ms in word-initial positions, followed by high vowel sounds.

The words “gift”, “geese”, and “geek” were chosen to measure the VOT value of the stop sound [g] in word-initial positions of C-V-C patterned words.

The mean VOT value of the [g] in “gift” in the NNS pronunciation was found to be 86.7 ms (SD=29.6). In “geese”, it was found to be 80.1 ms (SD=38.9), and in “geek”, it was 82.2 ms (SD=51.2).

The mean VOT values for the exact words in NSs’ pronunciation were found to be 50.5 ms (SD=29.7), 67 ms (SD=55.5), and 58.7 ms (SD=32), respectively, which is not conformant with previous findings reported by Lisker & Abramson (1964), Klatt (1975), and Caruso & Burton (1987).

The mean VOT value of “geese” ranged from 162 ms to 29 ms and from 15 ms to 178 ms for NNSs and NSs, respectively. Interestingly, when a t-test was conducted with the mean VOT values produced by NNs and NNSs, it was established that there is no statistically significant difference in the mean VOT value ($t(18) = 0.611$; $p=.549$, two-tailed).

VOT VALUE OF STOPS IN LOW VOWEL [a]

VOT OF [p] AND [b]

Three experimental words starting with stop sound [p], which have a C-V-C sound pattern, where the medial vowel sounds are [a:] and [a] were selected. The words are “pot”, “pomp”, and “pocket”.

Most of the NNSs produced shorter VOTs compared to NSs in “pot”. The longest VOT produced by NNSs was found in subject 1, which was equal to 67 ms. The longest VOT among the NSs was in subject 2, which was equal to 157 ms. The mean VOT value for NNSs come to 24.5 ms (SD=16.2), while it came to 55.6 ms (SD=37.1) for NSs. The t-test showed a significant difference in VOT production in the two groups of subjects (NNSs and NSs, $p=.026$).

The mean VOT produced by NNSs 1, 8, and 9 was the same as the VOT value produced by the NSs 1, 4, and 7. The longest VOT was 136 ms produced by NSs, and the shortest VOT was 16 ms. The longest VOT produced by NNSs was 56 ms. Overall, NNSs tend to produce shorter VOT compared to NSs. The mean VOT value of the VOT produced by 10 NNSs accounted for 37.9 ms (SD=24.9), and the mean VOT of the NSs accounted for 66.3 ms (SD=33.2). Interestingly, a t-test has shown a statistically significant difference between the mean VOT produced by NNSs and NSs ($t(18) = -2.16$; $p=.044$, two-tailed).

In the word “pocket”, NNSs produced the mean VOT of 28.2 ms (SD=12.5) while the NSs produced the mean VOT value, which is equal to 66.5 ms (SD=26.5). The t-test with the given word has shown a significant difference in the mean VOT ($t(18) = -4.118$; $p=.001$).

In the NS VOT values, the research findings are conformant with the findings of Lisker & Abramson (1964) and Caruso & Burton (1987). However, findings also show that the sound [p] is relatively problematic for NNSs in VOT value.

For the case of [b] sounds in word-initial position, the selected words for the comparison and contrast of VOT length were: “body”, “bomb”, and “boggle”.

The mean VOT length of the sound [b] in “body” for the NNSs was found to be 77.1 ms (SD=48.1), whereas, for NSs, it came to 21.2 ms (SD=25).

NNSs tend to produce longer VOTs compared to all NSs. For example, the longest VOT was found in NNS subject 8 at 142 ms, and the shortest VOT value was found in subject 5 at 9 ms. Among NSs, the longest VOT value was produced by subject 4 at 73 ms.

Most NSs tend to produce the sound [b] in a word-initial position with zero VOT. It was also found that compared to NSs, NNSs tend to produce the sound [b] in a word-initial position with a positive VOT.

In the case of “bomb”, the results were found to be similar to “body”. The mean length of the VOT in NNSs was found to be 97.1 ms (SD=36.3). Compared to NNSs, NSs produced the mean VOT value of 22 ms (SD=30.7), which is significantly shorter than NNSs’ VOT.

The mean VOT value of the sound [b] in “boggle” was found to be 101.4 ms (SD=41.6) for NNSs and 25.3 ms (SD=31) for NSs. It can be seen that there is a large discrepancy in the mean VOT values of the two. The results of a t-test show that there is a significant difference in the mean VOT values.

In the studies of Caruso & Burton (1987) and Macleod & Stoel-Gammon (2005), the VOT value of the sound [b] was reported to be 19.7 ms and 19.8 ms, which makes the findings of the given research conformant with them. Regarding Uzbek speakers, the VOT values of the sound [b] in the given experimental words ranged from 77 to 100 ms.

VOT OF [t] AND [d]

A comparison and contrast of the VOT length in NSs and NNSs in stop sounds [t] and [d] “tom”, “tarn”, “target”, “dark”, “dahl”, and “doll” were chosen.

The mean VOT value of [t] in “tom” was found to be 43.5 ms (SD=21.9) in NNSs. In NSs, the mean VOT value was found to be 95.9 ms (SD=42.6).

The VOT values of NSs are significantly longer than NNSs’ VOT. For all NNSs, the VOT value was below 80 ms, whereas for all of the NSs VOT value is longer than 50 ms. In the case of “tarn”, the mean VOT value for the NNSs was found to be 51.5 ms (SD=25.6) and was 106.5 ms (SD=52) for NSs.

When the means for the word-initial stop [t] in “target” were calculated, it was found that the mean for the NNSs was 39.7 ms (SD=19). For the NSs, it was found to be 87.5 ms (SD=31.9).

The study findings regarding NSs follow the findings reported by Kessinger and Blumstein (1997) and Macleod and Stoel-Gammon (2005). The VOT value of the sound [t] in their studies was reported to be 87.9 ms and 100 ms respectively. As for the NNSs, they tend to produce the sound [t] in an Uzbek manner, which shortens the VOT value to the value reported by Mirtozhiev (2013) in Uzbek sound [t].

In the case of “dark”, where the word-initial sound is stop sound [d], the calculations of the mean VOT values showed the following results. The mean VOT value for the NNSs was found to be 93.5 ms (SD=37), and for the NSs, it was found to be 39.2 ms (SD=41).

Among the NSs, there was no VOT which is longer than 113 ms. Most of the subjects in the NS group produced mean VOTs of below 20 ms, which is conformant with the findings of Lisker and Abramson (1964), Klatt (1975), Caruso and Burton (1987), Kessinger and Blumstein (1997) and Macleod and Stoel-Gammon (2005).

In “dahl”, the VOT mean value for the sound [d] was found to be 73 ms (SD=45.3) for NNSs and 49.7 ms (SD=47.1) for NSs.

The findings of the research in terms of the given word differ strongly from the findings reported by Lisker and Abramson (1964); Klatt (1975); Caruso and Burton (1987); Kessinger and Blumstein (1997); and Macleod and Stoel-Gammon (2005). However, closer examination of individual NSs shows that there are subjects (2 out of 10) whose VOT value was in line with the VOT reported by previous researchers.

A t-test was conducted to compare the mean VOT value for NNSs and NSs. The t-test showed no significant difference in the mean VOT value for NNSs (M = 73, SD = 45) and NSs

($M = 49.7$, $SD = 47$); $t(18) = 1.1$; $p = 0.275$, two-tailed). Overall, according to the statistical data analysis results, there is no difference in the mean VOT value of [d] in “dahl”.

As for “doll”, the mean VOT value was found to be 106.3ms ($SD=28.4$) for NNSs and 25.5 ms ($SD=24.9$) for NSs. Furthermore, a t-test showed a statistically significant difference in the mean VOT value ($t(18) = 6.7$; $p = .000$, two-tailed).

In the word “doll”, it can be seen that the VOT value of the NSs slightly differs from the VOT reported by Caruso & Burton (1987); Kessinger & Blumstein (1997); and Macleod & Stoel-Gammon (2005).

VOT OF [k] AND [g]

Three experimental words starting with stop consonant [k], which have a C-V-C syllabic pattern, where the medial vowel sound is /a/ were selected. The words are “Kahn”, “collar”, and “condo”.

In most cases, NNSs’ VOT was found to be almost the same as NSs’ VOT. However, the most significant discrepancy was observed in subjects 2, 3, 5, and 10 in both subject groups.

The mean VOT value for NNSs was 73.3 ms ($SD=47.8$), while it came to 96.7 ms ($SD=34.4$) for NSs. T-test results have shown that there is no statistically significant difference in the mean VOT of the sound [k] in native and NNSs’ production of the word “kahn” ($t(18) = -1.25$; $p = .225$, two-tailed).

In “collar”, the mean VOT of [k] was found to be 71.1 ms ($SD=38.4$) for NNSs and 107.8 ms ($SD=49.8$) for NSs. A t-test was conducted to compare the mean VOT value for NNSs and NSs. The t-test showed no statistically significant difference in the mean VOT value for NNSs ($M = 71.1$, $SD = 38$) and NSs ($M = 107.8$, $SD = 49.8$); $t(18) = -1.8$; $p = 0.082$, two-tailed).

The mean VOT of [k] in “condo” was found to be 56.6 ms ($SD=13.4$) for NNSs and 87.2 ms ($SD=35.5$) for NSs. Thus, in the given experimental word, NSs tend to produce longer VOTs compared to NNSs.

Overall, the findings of the research in terms of the NSs are conformant with the studies carried out by Lisker and Abramson (1964), Klatt (1975), and Caruso and Burton (1987), who reported the VOT value of [k] to be 80, 70, and 74.8 ms, respectively. Furthermore, NNSs produced the same VOT value reported by the studies mentioned above, confirming that for NNSs, the stop sound [k] is not problematic.

Three experimental words starting with stop sound [g], with C-V-C syllabic pattern, where the medial vowel sounds are [a:] and [a] were selected. The words are “garbage”, “garden”, and “garlic”.

In “garbage”, the mean VOT value of [g] was found to be 61.2 ms ($SD=48.7$) in NNSs. On the other hand, the VOT value of the NSs was found to be 27.1 ms ($SD=11.3$), which follows the findings of Klatt (1975).

In “garden”, the mean VOT value was found to be 76.3 ms ($SD=38.4$) for NNSs and 28.9 ms ($SD=10.1$) for NSs. In addition, the results of a t-test indicated a statistically significant difference in the mean VOT values of the [g] in “garden” pronounced by NNSs and NSs.

The results of [g] VOT mean values in “garlic” produced by NNSs and NSs showed that the mean VOT for NNSs was found to be 73.9 ms ($SD=37.3$), whereas, in the NSs, it was found to be 38.3 ms ($SD=24.8$). The results indicate that the mean VOT of [g] in “garlic” was twice as long in NNSs compared to NSs. A t-test showed a statistically significant difference between the mean VOT of NS and NNS of [g] in “garlic”.

The research findings regarding the NSs’ VOT values align with the findings reported by Klatt (1975) and Caruso and Burton (1987). In all three experimental words with the word-initial sound [g] with the preceding sound [a], NNSs’ mean VOT was longer than NSs. This

can probably be explained by the fact that in the Uzbek language, the VOT value [g] is considerably longer (150 ms), as reported by Mirtozhiev (2013), than the VOT value of the English sound [g].

Besides, strong mother tongue interference is observed in the production of the experimental words, which causes a significant difference in VOT values.

VOT VALUE OF WORD-INITIAL STOPS IN HIGH VOWEL [u]

VOT OF [p] AND [b]

Three experimental words starting with stop sound [p], which have a C-V-C syllabic pattern, where the medial vowel is [u], were selected. The words are “pool”, “poodle”, and “pooh”.

In “pool”, the mean VOT value was found to be 61 ms (SD=27.3) for NNSs and 100.3 ms (SD=50.2) for NSs.

In regards to “poodle”, the mean VOT value for the NNSs was found to be 38.3 ms (SD=14.9), whereas, for the NSs, it was found to be 82.1 ms (SD=26.2).

In “pooh”, the VOT value of the sound [p] was found to be 44 ms (SD=18.7) for NNSs and 78.3 ms (SD=19.6) for NSs. Among the NNSs, the longest mean VOT value was 76 ms. On the other hand, the shortest mean VOT was 54 ms for NSs and 21 ms for NNSs. Mirtozhiev (2013) reported that the VOT value of the sound [p] ranges from 40 to 60 ms, which is also reflected in the pronunciation of the English sounds.

The VOT of [b] followed by high vowel [u] was tested with the experimental words “boo”, “booby”, and “boom”.

In “boo”, the mean VOT was found to be 104.2 ms (SD=54.2) for NNSs, and 51.2 ms (SD=66.3) for NSs.

NSs tend to produce shorter VOTs compared to NNSs. The longest mean VOT found among the NNS was 183 ms, and the longest mean VOT for the NSs was equal to 207 ms. In “booby”, the mean VOT of the [b] was found to be 56.6 ms (SD=34.7) for NNSs and 24.3 ms (SD=24.4) for NSs.

As for the NSs, most of the VOT values are conformant with the findings of Caruso & Burton (1987) and Macleod & Stoel-Gammon (2005), who found VOT values of 19.7 ms and 19.8 ms for [b].

As regards “boom”, the mean VOT value of the word-initial stop [b] was found to be 80.2 ms (SD=50.4) for NNSs and 33 ms (SD=40.3) for NSs. Overall, the mean VOT value of [b] in all three experimental words, namely, “boo”, “booby”, and “boom”, was longer in NNSs. In the Uzbek language, the VOT value of [b] ranges from 150 to 200 ms (Mirtozhiev 2013), which confirms that when NNSs pronounced the sound [b], they faced negative interference of their native language.

VOT OF [t] AND [d]

Three experimental words starting with stop sound [t], which have a C-V-C syllabic pattern, where the vowel sound is [u], were selected. The words are “tomb”, “tool”, and “tooth”.

In “tomb”, the mean VOT of [t] was found to be 31.3 ms (SD=17.7) for NNSs. For the NSs, the VOT value was found to be 93.5 ms (SD=24.3), which is in line with the findings of Kessinger & Blumstein (1997) and Macleod & Stoel-Gammon (2005). Overall, NNSs’ mean VOT of [t] in “tomb” was not longer than 60 ms, while among the NSs, it was not less than 60 ms. T-test results indicate a statistically significant difference in the mean VOT of the NNSs and NSs ($p=0.000$; two-tailed). NS produced a VOT value equal to 0.129 ms, whereas NNS produced a VOT value of 0.35 ms, which is considerably shorter than the NS VOT value. In

addition, NS produced the stop [t] with strong frication at the beginning of VOT, whereas frication in NNS pronunciation of [t] was considerably weaker.

The high discrepancy in the pronunciation of the sound [t] in “tomb” can be because most NNSs mispronounced the word. As a result, they tend to transfer high vowel sound [u] into [o], which negatively influences the VOT value.

Next, the results of VOT analysis and mean calculations revealed that the mean VOT of the sound [t] in “tool” was 62.4 ms (SD=25.1) for NNSs and 107.9 ms (SD=46.1) for NSs. The mean VOT of the NNSs ranges from 14 ms to 108 ms. Among the NSs, the VOT mean ranges from 72 ms to 222 ms. NNSs’ VOT values were close to the VOT values reported by Klatt (1975), while the VOT values of the NSs was close to the VOT values reported by Kessinger & Blumstein (1997).

Overall, in “tool”, NNSs’ mean VOTs were significantly shorter than those of native speaker subjects.

In “tooth”, the mean VOT of [t] was found to be 66.2 ms (SD=11.8) for NNSs and 90 ms (SD=33.7) for NSs. Thus, the research findings in terms of the NSs are conformant with the findings reported by Kessinger & Blumstein (1997) and Macleod & Stoel-Gammon (2005). Interestingly, NNSs’ VOT values in “tooth” were close to the VOT value [t] reported by Klatt (1975).

The study results showed that NSs produced a longer mean VOT of [t] than NNSs in “tooth” in most cases.

In case [d], three experimental words, “doom”, “doomster”, and “doozy” with sound [u], were chosen. In “doom”, the mean VOT of [d] was found to be 80.6 ms (SD=42.2) for NNSs. For the NSs, the mean VOT was 40.8 ms (SD=38.1), which is not conformant with the results reported in the literature. Although the mean VOT in the whole group of NSs differs significantly from the VOT value reported in the literature, individual examination of the VOT values shows that the values are conformant with the ones reported by previous studies (Klatt 1975; Caruso & Burton 1987; Kessinger & Blumstein 1997; Macleod & Stoel-Gammon 2005).

The results indicate that the NNSs’ mean VOT of [d] in “doom” was longer than the NSs’ mean VOT. The low standard deviation (3.2) of the mean values in the NSs group can be explained by the mean VOT value among NSs was not higher than 0.24 ms, whereas among the NNSs, the highest VOT value was 0.143 ms and the lowest was 0.026 ms.

The word “doozy” was the following word in the list: to compare the mean VOT of the sound [d] in NS and NNS pronunciation. The results showed that the mean VOT of the sound [d] in “doozy” was found to be 94.9 ms (SD=63.8) for NNSs. However, for NSs, it was found to be 23.9 ms (SD=7.5) which does not match or slightly differs from the VOT values reported by Caruso & Burton (1987) and Kessinger & Blumstein (1997).

VOT OF THE [k] AND [g]

Three experimental words, “cook”, “cool”, and “coon”, with the word-initial stop sound [k] and vowel [u], were chosen for the experiment.

In “cook”, the mean VOT of [k] was found to be 69.7 ms (SD=20.3) for NNSs and 79.1 ms (SD=31.9) for NSs. Thus, both NNSs and NSs’ VOT values follow the findings reported in the literature (Klatt 1975; Caruso & Burton 1987).

A t-test was conducted to compare the mean VOT value for NNSs and NSs. The t-test showed no statistically significant difference in the mean VOT value for NNSs ($M = 69.7$, $SD = 20$) and NSs ($M = 79.1$, $SD = 32$); $t(18) = -.785$; $p = 0.443$, two-tailed).

Although NS produced VOT with less frication than NNS, the VOT of NNS (0.065 ms) was close to NS’s VOT value (0.063).

As for “cool”, the mean VOT of [k] was found to be 80.7 ms (SD=20) for NNSs, and for the NSs, the mean VOT of [k] was found to be 119.8 ms (SD=35.7). Thus, the NS’s VOT value differs considerably from that reported if in the literature.

The results of the mean VOT calculation revealed almost the same results as in “coon”. The mean VOT of [k] in “coon” was 75 ms (SD=14.3) for NNSs and 106.5 ms (SD=27.6) for NSs. Interestingly, the mean VOT of the NNSs is in line with the findings reported by Caruso & Burton (1987), while the NSs’ VOT differ significantly from the ones reported in previous studies.

To compare and contrast the mean VOT of the sound [g] in words where the vowel is [u], three words were chosen: “goof”, “google”, and “goose”.

The mean VOT values of [g] in the three words were 80 ms, 92 ms and 102 ms for NNSs. In the NS group, the mean VOT values of [g] in “goof”, “google”, and “goose” were found to be 28 ms, 30 ms and 51 ms, respectively.

In “google”, the longest mean VOT was 154 ms in the NNSs’, and the shortest mean VOT was 33 ms. In contrast, almost all NSs produced a significantly shorter mean VOT, with 51 ms being the longest and 10 ms being the shortest mean VOT. Two subjects out of 10 in the NNSs’ group produced a mean VOT that was almost similar in length to NSs.

The NSs’ mean VOT of [g] in “goose” was considerably shorter, 178 ms being the longest and 16 ms being the shortest mean VOT among them all. Only one subject among NNSs produced a mean VOT which is close to NSs’ mean VOT. Findings of the given research regarding the VOT values of the sound [g] in “goof” and “google” match the findings reported in the studies carried out by Klatt (1975) and Caruso & Burton (1987). However, the VOT value of [g] in “goose” differs significantly from the one reported in the literature.

According to the research results, NNSs’ mean VOT values in the production of stops were statistically similar to NS pronunciation in 9 experimental words out of 54. These include “tooth”, “dahl”, “key”, “keep”, “kahn”, “collar”, “cook”, “geek”, and “geese” (see Figure 1).

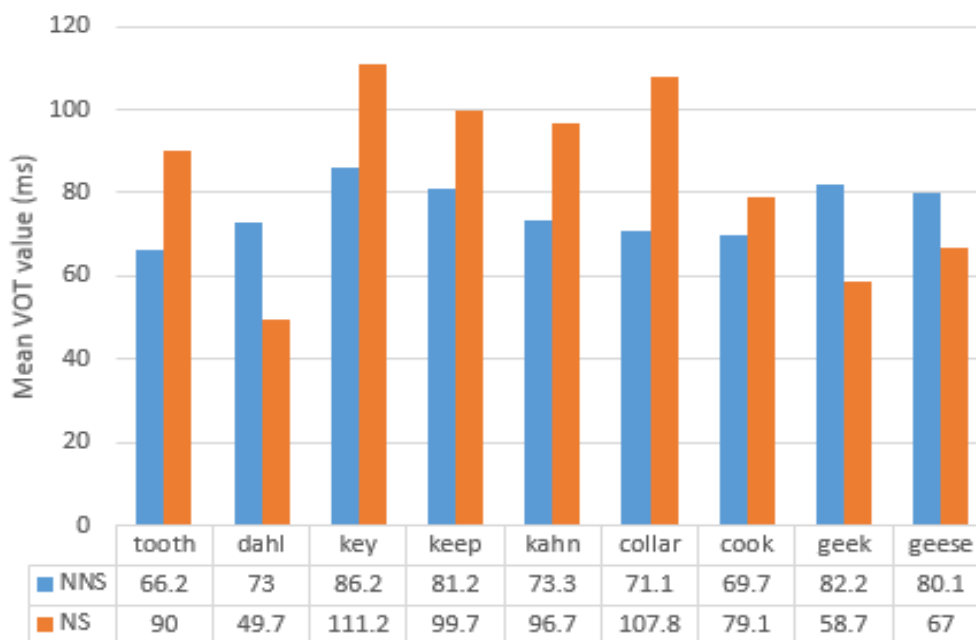


FIGURE 1. Words With No Statistically Significant Difference in Vot Values

In a majority (45) of the experimental words, the NNSs’ mean VOT of the voiced stop consonants /b/, /g/, /d/ and voiceless stop consonants /p/, /k/, /t/ differed significantly from NSs’ production of similar sounds in the same experimental words.

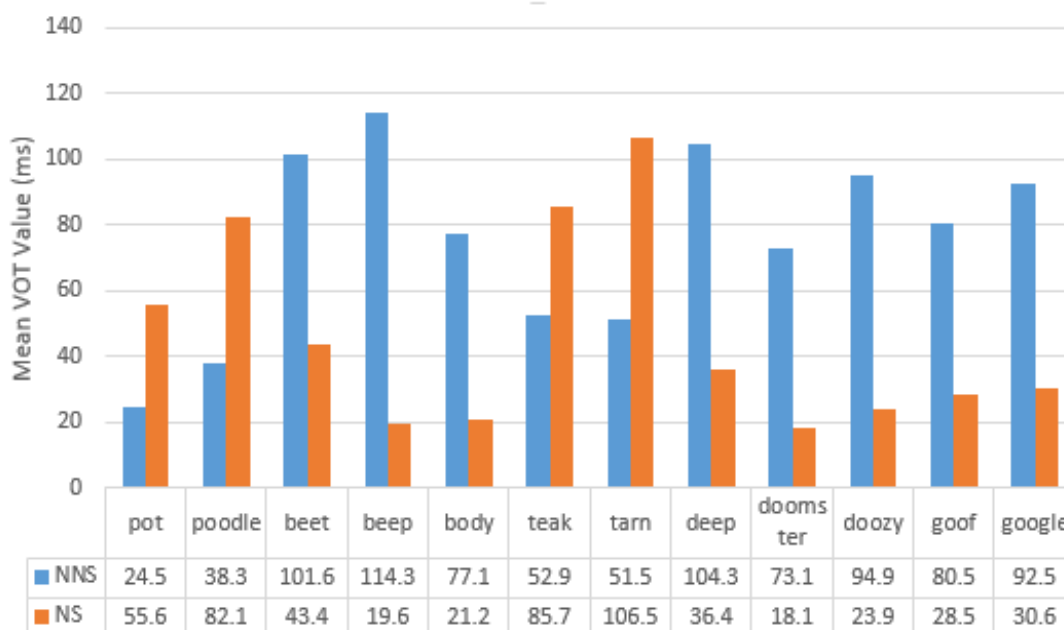


FIGURE 2. Words with Statistically Significant Difference in Vot Values

CONCLUSION

The research findings indicate that Uzbek speakers' production of English stop consonants differs significantly in most of the tested experimental words in terms of the VOT.

The experimental words with word-initial stop consonants were divided into categories according to the preceding vowels and their nature. Findings suggest that in some words where word-initial stop consonants are followed by a high vowel [i:], Uzbek speakers produce almost the same mean VOT compared to American English NSs.

It can be concluded that Uzbek speakers' production of English stop consonants in word-initial positions, where word-medial sounds are vowels of different qualities, differs significantly from native English speakers' production of similar stop consonants.

Overall, the given research rejects Yusupov's (2013) claim that the Uzbek language possesses consonants that are entirely similar to English consonants. This could be explained by the fact that there have been no comparative studies conducted on periodic characteristics of pronouncing consonant sounds. Furthermore, no studies have been conducted comparing the VOT values of English and Uzbek stop consonants.

Findings of the given research suggest that, first, teaching English consonants to Uzbek English speakers should consider periodic measurement of the VOT values produced by these English learners. This would help determine how succeeding vowel sounds impact the production of word-initial consonants and declare specific pronunciation instruction strategies that may modify Uzbek learners' production of plosive consonants and succeeding vowels. Second, establishing a periodicity of consonants will also help Uzbek English learners know to what extent to shorten or lengthen the VOT in their production of stop consonants. Next, the findings of the study will contribute to the body of knowledge on the instrumental study of English sounds produced by Uzbek learners of English, which will serve as a basis for carrying out further studies on VOT analysis of the Uzbek language sound system and comparative phonology of the Uzbek and English languages. Finally, taking into account that VOT may vary in spontaneous speech (Yao 2009), it is suggested that further studies be conducted with the spontaneous speech of Uzbek speakers of English.

REFERENCES

- Abduazizov, A. (1981). *Elements of general and comparative typological phonology*. Fan.
- Abduazizov, A. (1992). *Phonology and morphophonology of the Uzbek language*. Uqituvchi.
- Allen, J., Joanne, L. & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, 113(1), 544-552.
- Allen, J. & Miller, J. (2004). Listener sensitivity to individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, 115(6), 3171-3183.
- Bakieva, G. & Rustamova, A. (2020). Exploring English pronunciation teaching in Uzbekistan. *International Journal of Scientific & Technology Research*, 9(1), 1208-1219.
- Bilá, M. & Eddy, E. (2012). *English phonetics and phonology for Slovak students*. Filozofická fakulta Presovskej univerzity.
- Boersma, P. & Weenink, D. (2019). Praat: Doing phonetics by computer (Version 6.0. 52) [Windows]. <http://www.praat.org>
- Caruso, A. & Burton, E. (1987). Temporal acoustic measures of dysarthria associated with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, 30(1), 80-87.
- Fowler C. A., Sramko, V., Ostrya, D. J., Rowland, S. A., Halle, P. (2006). Voice onset time of bilingual English and French-speaking Canadians. *The Journal of the Acoustical Society of America*, 119(5), 3422-3422.
- Higgins, M., Netsell, R. & Schulte, L. (1998). Vowel-related differences in laryngeal articulatory and phonatory function. *Journal of Speech, Language, and Hearing Research*, 41(4), 712-724.
- Kessinger, R. & Blumstein, S. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, 25(2), 143-168.
- Kiliç, M. (2018). Voice onset time as an indicator of interlanguage: Evidence from Turkish-accented English. *Dil ve Dili Bilimi Çalışmaları Dergisi*, 14(4), 272-284.
- Klatt, D. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, 18(4), 686-706.
- Lisker, L. & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), 384-422.
- MacLeod, A. & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders*, 3(2), 118-127.
- McMahon, A. (2002). *An introduction to English phonology*. Edinburgh University Press.
- Mirtozhiev, M. (2013). *Phonetics of the Uzbek language*. Uqituvchi.
- Munro, M. & Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, 45(1), 73-97.
- Mustafa, A. (2019). Perception of American English allophonic variant /r/ by Arabic Speakers in an EFL Context. *3L: Language, Linguistics, Literature®*, 25(2), 127-142.
- Nabieva, D. (2005). *Reflection of dialectical generality and adherence in different layers of the Uzbek language*. Sharq.
- Nurmonov, A. (1990). *Phonology and morphophonology of the Uzbek language*. Uqituvchi.
- Nurmonova, D. (1998). *Phonological resistance and their moderation*. Sharq.
- Olmstead, A., Viswanathan, N., Aivar, P. & Manuel, S. (2013). Comparison of native and non-native phone imitation by English and Spanish speakers. *Frontiers in Psychology*. Retrieved May 4, 2021 from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3722886/>
- Olson, E. (2017). Voice onset time in Arabic and English Stop Consonants. A Senior Honors Thesis, Faculty of the University of Utah, USA.
- Shoabdurahmonov, Sh., Askarova, M., Hojiev, A. & Rasulov, I. (1980). *Modern Uzbek literary language*. Uqituvchi.
- Schoonmaker-Gates, E. (2015). On Voice-Onset Time as a Cue to foreign accent in Spanish: Native and non-native perceptions. *Hispania*, 98(4), 779-791. <http://www.jstor.org/stable/24572246>
- Taehong, C. & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of phonetics*, 27(2), 207-229.
- Tursunov, U., Urinboev, B. & Aliev, A. (1995). *History of the Uzbek literary language*. Uqituvchi.
- Yao, Y. (2009). Understanding VOT variation in spontaneous speech. *UC Berkeley PhonLab Annual Report 5.5*. <https://doi.org/10.5070/P76dd1x6cs>
- Yusupov, U. (2013). *Contrastive linguistics of the English and Uzbek languages*. Tashkent: Uqituvchi.

APPENDIX

LIST OF EXPERIMENTAL WORDS USED IN THE STUDY

/i/, /i:/

p/b: (1) a. peat, peace, peek

b. beet, beep, beak

t/d: (2) a. teak, tease, teat

b. deep, deed, dean

k/g: (3) a. keen, keep, key

b. gift, geese, geek

/a/, /a:/

p/b: (1) a. pot, pomp, pocket

b. body, bomb, boggle

t/d: (2) a. Tom, tarn, target

b. dark, dahl, doll

k/g: (1) a. Kahn, collar, condo

b. garbage, garden, garlic

/u/, /u:/

p/b: (1) a. pool, poodle, pooh

b. boo, booby, boom

t/d: (2) a. tomb, tool, tooth

b. doom, doomster, doozy

k/g: (3) a. cook, cool, coon

b. goof, Google, goose