

Voice onset times for Turkish stop consonants

Fatih Ögüt^{a,*}, Mehmet Akif Kiliç^b, Erkan Zeki Engin^c, Raşit Midilli^a

^a *Ege University, Medical Faculty, ENT Department, 35100 Izmir, Turkey*

^b *Sütçü İmam University, Medical Faculty, ENT Department, Kahramanmaraş, Turkey*

^c *Ege University, Engineering Faculty, Electric and Electronics Department, Izmir, Turkey*

Received 28 January 2005; received in revised form 16 October 2005; accepted 10 February 2006

Abstract

In this study, we aimed to determine the average VOT (voice onset time) values of the Turkish stop consonants by using 30 volunteers (15 female and 15 male). For this aim, we measured the VOT values of the six Turkish stops (i.e., /p/, /b/, /t/, /d/, /k/ and /g/), which were uttered by 30 subjects in three times, on wideband spectrograms. At the result of this study, the average VOT values of /p/, /b/, /t/, /d/, /k/ and /g/ were found to be 41, –66, 50, –53, 69, and –10 ms, respectively. © 2006 Published by Elsevier B.V.

Keywords: Articulation; Consonant; Acoustics; Speech; Stop consonants; Voice onset time

1. Introduction

The voice onset time (VOT) is a temporal acoustic parameter defined as the time between the release of the oral constriction for plosive production and the onset of vocal fold vibrations (Lisker and Abramson, 1964). VOT is known to be the most reliable acoustic cue for the distinction between voiced and voiceless stops and this temporal characteristic of stop consonant reflects the complex timing of supralaryngeal–laryngeal coordination (Abramson, 1977).

The stop consonants in Turkish may be classified into three groups according to the place of their articulation: bilabials (/p/, /b/), dentals (/t/, /d/), and velars (/k/, /g/). For each group, it has two different types according to the manner of articulation:

voiced and voiceless. The /t/ and /d/ consonants in Turkish are generally dental and less frequently alveolar unlike English. The /k/ and /g/ consonants in Turkish are velar ([k] and [g]) in back vowel environment, and palatal ([ç] and [j]) in front vowel environment. Therefore, it can be said that they have two allophones: velar and palatal (Demircan, 1996; Kornfilt, 1997). The three voiceless stops are aspirated in stressed syllables, and may have very long VOT values.

Measurements of VOT before the release are stated as negative numbers and called “voicing lead”, while measurements of VOT after the release are stated as positive numbers and called “voicing lag” (MacKay, 1987). If release and voicing are simultaneous, VOT is zero.

Extensive cross-language studies (Lisker and Abramson, 1964, 1970, 1971) have demonstrated that three categories of stops emerge from the VOT continuum:

* Corresponding author. Tel./fax: +90 232 3903757.
E-mail address: fatih.ogut@ege.edu.tr (F. Ögüt).

- (1) “Voicing lead”: ranging from about –125 to –75 ms with a median value of –100 ms. Italian voiced stops are of this type.
- (2) “Short voicing lag”: ranging from 0 to +25 ms with a median value of +10 ms. Italian voiceless stops and English voiced stops are of this type.
- (3) “Long voicing lag”: ranging from +60 to +100 ms with a median value of +75 ms. English voiceless stops are of this type.

Factors that can affect the magnitude of VOT include physiological differences (such as age, lung volume), pathological status (hearing impairment, depression), and different linguistic tasks (speech task, speech rate, phoneme environment) (Auzou et al., 2000).

VOT is usually measured from wideband spectrograms according to the procedure recommended by Lisker and Abramson (1964). It is measured from the onset of the energy ‘burst’ corresponding to the release of an articulatory constriction to the first of the regularly spaced vertical striations of the vocal fold vibrations.

There is only little difference in measurements obtained from oscillogram and spectrograms (Petrosino et al., 1993). However, time-synchronized spectrographic and oscillographic displays are now easily obtained with digitized acoustic signals, providing a greater accuracy of measure (Sweeting and Baken, 1982; Davis, 1995). The measurements on wideband spectrograms and oscillograms are highly correlated and the combined use of them may reduce the frequency of errors due to a short span of pulsation (Auzou et al., 2000).

In some pathological voices, VOT cannot be measured. This can either be due to the impossibility to determine the location of the burst or the onset of regular striations on the spectrogram (Özsancak et al., 2001). The burst does not occur when a subject fails to achieve full closure in the production of stop constants. Similarly, in normal voices the vowel onset may be difficult to determine (Sweeting and Baken, 1982). Few researchers’ report the percentage of unmeasurable VOTs and the data available indicate that less than 4% of the productions cannot be measured (Hardcastle et al., 1985; Itoh et al., 1982; Sweeting and Baken, 1982).

Standard VOT values of the languages such as English, Spanish, and French, etc. were determined and for these languages, the effects of articulations (place and manner) were also investigated. How-

ever, in the literature, we have not encountered any research about the VOT values of the Turkish stop consonants. Therefore, we measured VOT values of the Turkish stops on the utterances of 30 volunteers and determined standard VOT values of them.

2. Materials and methods

2.1. Subjects

Our subjects were 15 male and 15 female monolingual native Turkish speaking volunteers from Aegean region but using a general Turkish dialect. Their ages ranged from 18 to 28 years (Mean: 22). All subjects underwent phoniatric and audiologic examination to approve that they had no speech and hearing problem.

2.2. Speech Stimuli

Each speaker uttered the following isolated syllables (the six Turkish stops, i.e., /p/, /t/, /k/, /b/, /d/, /g/ combined with the eight Turkish vowels, i.e., /a/, /ɛ/, /u/, /i/, /ɔ/, /œ/, /u/, /y/) in three times in the following single syllable words in a carrier sentence: [pas], [pɛc], [puu], [pil], [pɔz], [pætf], [pul], [pyf]; [bal], [ben], [buuk], [biu], [bof], [bœn], [buz], [byc]; [taf], [tɛc], [tuup], [tiz], [tɔp], [tœz], [tuz], [tym]; [dam], [dɛv], [duuf], [dic], [dɔl], [dœn], [duu], [dyf]; [kas], [cɛl], [kuuf], [cim], [kɔtf], [cœu], [kuf], [cyl]; [gɔf], [jɛz], [guuk], [jit], [gɔl], [jœtf], [gut], [jyz].

2.3. Recording procedure

Subjects’ utterances were recorded in a sound-treated booth (IAC) by using a Seinheiser K6 omnidirectional microphone and a PC with Sound Blaster Live sound card. The microphonemouth distance was constant at 15 cm. Each syllable was saved at 11 025 Hz sampling rate and at 16-bit resolution. For recording and analysis Multi-Speech (Kay Elemetrics, Model 3700, version: 2.4) program was used.

The measurements were done on wideband spectrograms according to the procedure recommended by Lisker and Abramson (1964). The recorded signals were pre-emphasized (0.8 pre-emphasis level) in order to emphasize the high frequency components, which normally contain much less energy than the lower frequency components. The measurement on wideband spectrograms (215 Hz,

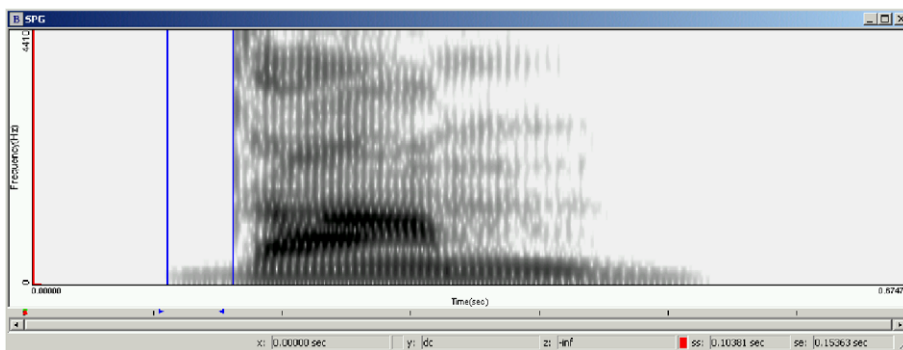


Fig. 1. Wideband spectrogram of the syllable [dom] in a normal subject (negative VOT value).

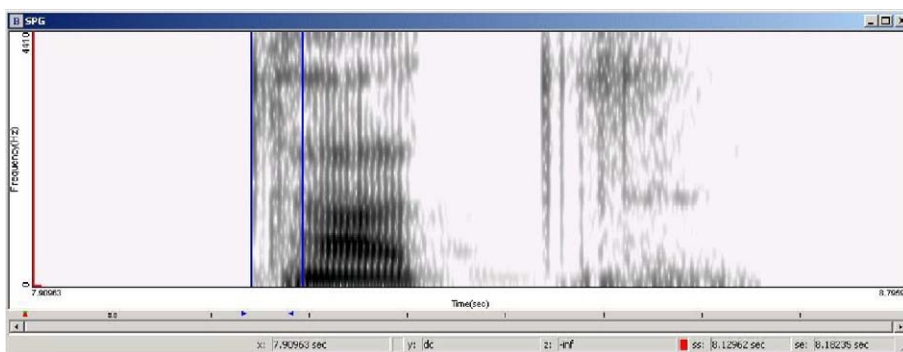


Fig. 2. Wideband spectrogram of the syllable [taɲ] in a normal subject (positive VOT value).

Blackman window) was done in the following way: For the stops having negative VOT values, VOT was measured from the onset of low-amplitude periodic voicing to the onset of aperiodic burst activity, usually signaled by a sharp peak of somewhat higher amplitude. This duration is expressed as a negative number (see Fig. 1). For the stops having positive VOT values, they were measured from the first indication of initial aperiodic activity to the regular striation, and thus is a positive value (see Fig. 2).

2.4. Statistical analysis

We used *statistical package for social sciences 11.0 (SPSS 11.0)* for all statistical analyses. A statistical significant level of $p = 0.05$ was used for all of the tests.

3. Results

The mean rate of measurable VOT value of utterances (3 repetitions \times 30 subjects' \times 6 stop con-

stants \times 8 vowels) was 95.4% because of the impossibility to determine the vowel onset.

Mean of VOT values for the six stop consonants combined with the eight vowels of female and male were separately given in Table 1. The effect of vowels on VOT values examined by using one-way ANOVA (analysis of variance) test and the p -values were found to be 0.892 for female and 0.977 for male. Thus, the vowel has no significant effect on VOT values.

To determine the average VOT values of the Turkish stop consonants, we used the average over the vowels for each of the stop consonant and the mean and standard deviation (SD) of these values for female and male were separately given in Table 2. One-way ANOVA test was used for determining the effects of sex and the p -value was found to be 0.499 which meant that the sex of the subjects' was not statistically significant. Therefore, we could use the averages of male and female VOT values (see Table 2) for the other statistical analyses.

The difference between average VOT values of voiced and voiceless was 107 ms for bilabial,

Table 1
Mean of VOT values in milliseconds for six stop consonants combined with the eight vowels of female (F) and male (M)

Vowel	Sex	Stop consonant					
		/p/	/b/	/t/	/d/	/k/	/g/
/a/	F	36	-51	46	-51	60	-8
	M	26	-74	43	-50	62	-8
/ε/	F	36	-36	42	-61	68	-14
	M	24	-84	36	-60	69	0
/u/	F	43	-66	60	-49	70	-13
	M	38	-82	54	-50	69	-13
/i/	F	49	-49	56	-48	75	-28
	M	42	-64	57	-53	78	-28
/ɔ/	F	49	-56	45	-66	62	8
	M	39	-79	42	-60	59	8
/æ/	F	41	-54	45	-44	66	-28
	M	34	-92	49	-50	69	-16
/u/	F	56	-61	58	-50	72	10
	M	45	-72	56	-56	74	-25
/y/	F	59	-63	55	-49	78	7
	M	41	-71	54	-54	78	-6

Table 2
Mean and standard deviation (SD) of VOT values in milliseconds for stop consonants of female, male, and average of them

	Female		Male		Both sexes	
	Mean (ms)	SD (ms)	Mean (ms)	SD (ms)	Mean (ms)	SD (ms)
/p/	45.3	11.7	36.2	21.6	40.7	17.6
/b/	-54.9	35.7	-77	41.5	-65.9	39.7
/t/	50.8	12.5	48.6	7.8	49.7	10.3
/d/	-51.8	37.4	-54	46.8	-52.9	41.6
/k/	68.9	13.7	69.8	7.2	69.4	10.8
/g/	-8.3	9.8	-11.2	31.4	-9.7	22.9

103 ms for dental, and 79 ms for velar stops. The effect of places of articulation examined by using one-way ANOVA test and the p -value was found smaller than 0.001 which meant that the place of articulation was statistically significant. After this analysis, to determine the differences between bilabials and dentals; bilabials and velars; dentals and velars, a Post hoc Tukey test was used and it was observed that bilabials and dentals differed from velars (p -values were <0.001 and 0.007, respectively) but not from each other ($p = 0.478$).

4. Discussion

Normal VOT values are different between languages like English, French, Italian, Spanish, and

Thai and the standardization must be made for all of the languages. The VOT studies have produced data for many languages, but have left out the analysis of Turkish. In the literature, we have not encountered any research about Turkish. Our paper may be a start to fill in the world VOT map.

It is also important to make the comparison of the VOT values measured for Turkish stops with those measured for other languages. We compared the VOT values regarding the techniques to measure the VOT, which have been used in different papers.

For four American English speakers the VOT values of /p/, /t/, /k/, /b/, /d/, /g/ were found to be 58, 70, 80, 1, 5, 21 ms, respectively (Lisker and Abramson, 1964). In the study of Macleod and Stoel-Gammon (2005), the VOT values were obtained from the waveform and verified with the spectrogram according to the procedure recommended by Lisker and Abramson (1964) for three Canadian English speakers and the mean VOT values of voiced stops /b/, /d/ and voiceless stops /p/, /t/ were 19.8 and 87.9 ms, respectively. The measurements were done on oscillograms for four English speakers and the VOT values of /p/, /t/, /b/, /d/ were found to be 85, 100, 15, 20 ms in the study of Kessinger and Blumstein (1997), respectively (the values were approximated from figures). For three subjects the VOT values of voiced stops /b/, /d/, /g/ were 11, 17, and 27 ms, respectively whereas the voiceless stops /p/, /t/, /k/ were 47, 65, and 70 ms, respectively (Klatt, 1975). For eight English speakers the VOT values of /p/, /t/, /k/, /b/, /d/, /g/ were found to be 62.5, 71.9, 74.8, 19.7, 21.4, 35.2 ms in the study of Caruso and Burton (1987) according to the procedure recommended by Klatt (1975).

For three Canadian French speakers the mean VOT values of voiced stops /b/, /d/ and voiceless stops /p/, /t/ were -99.3 and 37.4 ms, respectively (Macleod and Stoel-Gammon, 2005). For four French speakers the VOT values of /p/, /t/, /b/, /d/ were found to be 30, 35, -115, -100 ms, respectively (the values were approximated from figures) (Kessinger and Blumstein, 1997). The measurements were done on oscillograms for five participants and the VOT values of voiceless stops /p/, /t/, /k/ were 45, 51, and 72 ms, respectively whereas the voiced stops /b/, /d/, /g/ were -140, -142, and -146 ms, respectively (Ryalls et al., 1995).

In the study of Bortolini et al. (1995), measurement of VOT values were made from both acoustic waveform and wide-band spectrograms according

to the procedure recommended by Lisker and Abramson (1964) for seven subjects and the VOT values of voiceless stops /p/, /t/, /k/ were 11.3, 19.3, and 34.1 ms, respectively whereas the voiced stops /b/, /d/, /g/ were -73.7, -79.9, and -66.9 ms, respectively.

For two Puerto Rican Spanish speakers the VOT values of voiceless stops /p/, /t/, /k/ were 4, 9, and 29 ms, respectively whereas the voiced stops /b/, /d/, /g/ were -138, -110, and -108 ms, respectively (Lisker and Abramson, 1964). The measurements were done on oscillograms for thirty-two Castilian Spanish speakers and the VOT values of /p/, /t/, /k/, /b/, /d/, /g/ were found to be 13.1, 14, 26.5, -91.5, -91.6, -73.7 ms, respectively (Rosner et al., 2000).

For three Thai speakers the VOT values of /p/, /t/, /k/, /b/, /d/ were found to be 6, 9, 25, -97, -78 ms in the study of Lisker and Abramson (1964), respectively. For four Thai speakers the VOT values of /p/, /t/, /b/, /d/, were found to be 15, 13, -70, -65 ms, respectively (the values were approximated from figures) (Kessinger and Blumstein, 1997).

In Turkish language the voiceless plosives have positive VOT values as in all other languages. The VOT value of /k/ is the longest in all of the languages. In the voiced plosives, the VOT values are negatives except for English. The VOT values for /b/ and /d/ are close to each other in all of the languages. Also, the VOT value of /g/ is close to the VOT values of /b/ and /d/ in English, French, Italian, and Spanish when compared in each language, but it differs a lot in Turkish. In Turkish, we obtained positive VOT values of /g/ in 12 (40%) and negative values in 18 (60%) of our cases.

Three categories of stops emerge along VOT continuum from extensive cross-language studies. The voiceless stops have long voicing lag and the voiced stops have short voicing lag for English whereas the voiced stops have voicing lead for French. For Italian, Spanish, and Thai, the voiced and voiceless stops have voicing lead and short voicing lag, respectively. The voiced and voiceless stops can be classified as voicing lead and long voicing lag for Turkish, respectively.

The subjects may well elicit very special pronunciations when they are using the actual words in carrier sentences. We used meaningful Turkish in the study, not isolated logatome syllables, to find out more about VOT in our language.

VOT may show variances according to the following vowel. The results may be ambiguous or distorted if we try to compare VOT values variance of two languages according to the following vowel. For example, the postconsonantal vowel was /ɔ/ rather than /a/ when Castilian Spanish speakers began voicing later for /p/, /k/ and started it earlier for /b/, /g/ (Rosner et al., 2000). We have made the comparison of VOT values variance according to the vowel only in our controlled environment. We found that following vowel did not significantly affect voicing in the Turkish language.

We also found that sex has no main effect on VOT values. Therefore, we determined the average VOT value of both sexes as a unique value.

The manner and place of articulations affect the values of VOT, thus it can be used as a reliable acoustic parameter that indicates the phonologic contrast between voiced and voiceless stops. Also, VOT could be used for the distinction of bilabials and dentals from velars but not from each other.

In the future studies the effects of physiological differences (age), pathological status (speech and hearing problems), and different linguistic tasks (speech rate, accent, and bilinguals) to Turkish VOT values will be further investigated. For this aim, methodological and technical properties of VOT measurement have to be standardized for Turkish language. Therefore, we used the method recommended by Lisker and Abramson (1964) to determine the Turkish VOT values.

References

- Abramson, A.S., 1977. Laryngeal timing in consonant distinctions. *Phonetica* 34, 295–303.
- Auzou, P., Özşancak, C., Morris, R.J., Jan, M., Eustache, F., Hannequin, D., 2000. Voice onset time in aphasia, apraxia of speech and dysarthria: a review. *Clinical Linguistics and Phonetics* 14, 131–150.
- Bortolini, U., Zmarich, C., Fior, R., Bonifacio, S., 1995. Word-initial voicing in the production of stops in normal and preterm Italian infants. *International Journal of Pediatric Otorhinolaryngology* 31, 191–206.
- Caruso, A.J., Burton, E.K., 1987. Temporal acoustic measures of dysarthria associated with amyotrophic lateral sclerosis. *Journal of Speech and Hearing Research* 30, 80–87.
- Davis, K., 1995. Phonetic and phonological contrasts in the acquisition of voicing: Voice onset time production in Hindi and English. *Journal of Child Language* 22, 275–305.
- Demircan, O., 1996. *Türkçenin Sesdizimi*. Der, Istanbul.
- Hardcastle, W.J., Barry, R.A.M., Clark, C.J., 1985. Articulatory and voicing characteristics of adult dysarthric and verbal dyspraxic speakers: an instrumental study. *British Journal of Disorders of Communication* 20, 249–270.

- Itoh, M., Samnuma, S., Tatsumi, I., Murakami, S., Fukusako, Y., Suzuki, T., 1982. Voice onset time characteristics in apraxia of speech. *Brain and Language* 17, 193–210.
- Kessinger, R.H., Blumstein, S.E., 1997. Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics* 25, 143–168.
- Klatt, D.H., 1975. Voice onset time, frication and aspiration in word initial consonant clusters. *Journal of Speech and Hearing Research* 18, 686–705.
- Kornfilt, J., 1997. *Turkish*. Routledge, London.
- Lisker, L., Abramson, A.S., 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20, 384–422.
- Lisker, L., Abramson, A.S., 1970. The voicing dimension: some experiments in comparative phonetics. In: *Proceedings of the 6th International Congress of Phonetic Sciences*, 563–567.
- Lisker, L., Abramson, A.S., 1971. Distinctive features and laryngeal control. *Language* 47, 767–785.
- MacKay, Ian R.A., 1987. *Phonetics and the Science of Speech Production*. Little, Brown and Co, Boston.
- Macleod, A.A.N., Stoel-Gammon, C., 2005. Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders* 3 (2), 118–127.
- Özsancak, C., Auzou, P., Mary, J., Hannequin, D., 2001. Measurement of voice onset time in dysarthric patients: Methodological considerations. *Folia Phoniatica et Logopaedica* 53, 48–57.
- Petrosino, L., Colcord, R.D., Kurcz, K.B., Yonker, R.J., 1993. Voice onset time of velar stop productions in aged speakers. *Journal of Perceptual and Motor Skills* 76, 83–88.
- Rosner, B.S., Lopez-Bascuas, L.E., Garcia-Albea, J.E., Fahey, R.P., 2000. Voice-onset times for Castilian Spanish initial stops. *Journal of Phonetics* 28, 217–224.
- Ryalls, J., Provost, H., Arsenault, N., 1995. Voice onset time production in French speaking aphasics. *Journal of Communication Disorders* 28, 205–215.
- Sweeting, P.M., Baken, R.J., 1982. Voice onset time in a normal-aged population. *Journal of Speech and Hearing Research* 5, 129–134.