# AN ACOUSTIC STUDY OF GREEK VOICELESS STOPS

Katerina Nicolaidis[1], Anna Sfakianaki[2], George Vlahavas[1], George Kafentzis[2]

[1]Aristotle University of Thessaloniki, [2]University of Crete, Greece
knicol@enl.auth.gr; asfakianaki@csd.uoc.gr; gvlahavas@csd.auth.gr; kafentz@csd.uoc.gr

## ABSTRACT

The paper investigates acoustic properties of the Greek voiceless plosives /p, t, k/, including the palatal allophone [c], by examining absolute and relative VOT and closure duration, relative burst intensity and spectral moments. Variability due to place of articulation, vowel context, gender and age is examined. The speech material comprised $C_1VC_2V$ real words ($C_1$=/p, t, k/, V=/i, a/, $C_2$=dental/alveolar). Data from 12 adult speakers and 12 children (6 male and 6 female in each group) were analysed. Results showed that relative closure duration decreased and relative VOT duration increased in the order /p/, /t/, /k/ showing the anticipated inverse relationship reported in the literature. VOT was longer in the high vowel context for /t, k/. All spectral moments were significantly affected by place of articulation. Relative burst intensity was greater for the velar. Effects of gender and age were variable. Results are discussed in relation to theory and crosslinguistic evidence.

**Keywords**: stops, closure, VOT, duration, burst, spectral moments, amplitude, Greek

## 1. INTRODUCTION

The voiceless stops of Modern Greek comprise the bilabial /p/, the dental or dentoalveolar /t/ and the velar /k/. The last has a palatal allophone [c] in the environment of the front vowels /i, e/ or /i/ followed by another vowel within the same syllable [3]. Its articulation can vary with constriction in the palatal or alveolopalatal area as shown in [17]. Voiceless stops are unaspirated in Greek. There are several studies that have examined the durational properties of the closure and release phases of the Greek voiceless stops (for example [2], [14], [18]; see also review in [3]) but there are hardly any studies examining other important acoustic properties such as burst amplitude, spectral moments and formant transitions to the flanking vowel environments (see e.g. [4] for selected spectral measurements of the stop release and the formant frequencies of flanking vowels, and [1] for a crosslinguistic study).

An important question in the study of stop production has been the determination of the acoustic correlates of place of articulation distinctions. Closure and VOT duration, burst spectral shape, spectral moments, burst amplitude, and formant transitions have been shown to contribute to place of articulation differentiation. Typically, as the point of articulation moves back in the oral cavity, closure duration decreases and VOT duration increases. As such, previous studies have reported closure durations decreasing and VOT durations increasing in the order /p, t, k/ (e.g. [7], [9], [15]). Physiological/aerodynamic characteristics have been shown to account for such variation (see [7] for a review). In particular, closure durational differences have been explained on the basis of the timing of the intraoral pressure peak which occurs more quickly when the supraglottal cavity behind the constriction is smaller. VOT differences have been explained on the basis of factors such as the volume of the cavity behind and in front of the constriction, the speed of articulatory release (i.e. fastest for lips, fast for the tongue tip, slow for the tongue body), the extent of the articulatory contact area (i.e. more extended contact relates to slower release).

Burst spectral shape differences among stops of different places have been noted with labials having a diffuse falling pattern, alveolars a diffuse rising pattern and velars a compact pattern. Such differences have been related to cavity shape differences in front of the occlusion [12]. In addition, spectral moments analyses have shown that more retracted places of articulation, which result in longer front oral cavity, are characterised by lower mean, lower standard deviation, positive skew, and positive kurtosis, i.e. lower frequency energy, more compact burst spectrum and more peaked distribution. Previous research on English stops has also shown that the strongest burst energy is found for the alveolars and the weakest for the labials ([10], [28]). Finally, important information on stop place of articulation is carried by F2 formant transitions in CV syllables (e.g. [5]. [24], [26]).

Furthermore, several factors have been reported to affect stop production. Significant contextual effects on the first spectral moment (M1) for /t/ and /k/ were shown in [20] indicating anticipatory V-to-C effects. The same study reports more anticipatory effects on /k/ for children than adults suggesting greater influence of the upcoming vowel on the tongue dorsum for children. The study also reports age related differences in M1 with higher values for children than adults due to the former's smaller oral cavities. In addition, age and gender effects have been reported on VOT duration in a series of studies and have been accounted for on the basis of physiological or sociophonetic factors, e.g. [19], [21], [22].

The current study aims to examine several acoustic characteristics of the Greek voiceless stops, i.e. absolute and relative closure and VOT duration, relative burst intensity and spectral moments of the burst. It examines variability in these measures due to factors such as place of articulation, context, gender, and age. This is the first study that combines all these measures including a spectral moments analysis of the stop bursts, and investigates variability in stop production due to the aforementioned factors. To our knowledge, gender and age effects have not been studied before for all these measures for the Greek stops. As such, the results of this study are expected to increase our understanding of the contribution of different factors to variability in stop production providing insights for theory and applied areas such as clinical practice and technological applications.

## 2. METHODOLOGY

The speech material consisted of $C_1VC_2V$ real words with $C_1$=/p, t, k/, V=/i, a/ (mostly symmetrical sequences), $C_2$ a dental/alveolar consonant and stress on the first syllable, e.g. /'pata/ (step (v)), /'pini/ (s/he drinks), /'kili/ (hernia). Words were embedded in the carrier phrase ['leje __ pa'du] "say __ everywhere" and were repeated three times by each subject at a comfortable speaking rate. Data were recorded from 12 adult subjects (A) and from 12 children (C) aged 8-10 (6 male (M) and 6 female (F) participants in each group). All subjects had normal hearing and speech and spoke Standard Modern Greek with no detectable accent. All children and most adults were from Thessaloniki; three adults came from other major cities in mainland Greece but lived in Thessaloniki for many years and two were from Athens.

Data were recorded with a Sennheiser ME66/K6 microphone in a sound-treated booth (sampling rate at 44.1 KHz). They were automatically segmented and were all manually checked and corrected, where necessary, by the first two authors in PRAAT. The following measures were then calculated: absolute stop closure duration (from the end of the formant frequencies of the final vowel in ['leje] to the burst), absolute VOT duration (from burst onset to the onset of regular pulsing for the following vowel), relative burst intensity, and spectral moments. Relative closure and VOT durations were also calculated to control for speech rate, i.e. percentage of the duration from closure onset and burst onset respectively to the offset of the first vowel. To normalize inherent intensity differences among individual speakers, relative burst intensity was calculated by subtracting the RMS intensity obtained from a 30ms Hamming window placed at the start of the vowel following the stop from the RMS intensity of a 10ms Hamming window centered on stop burst onset (dB). A 0 value indicated that the first 30ms of the vowel and the burst intensity were the same.

Spectral moments (mean, standard deviation, skewness, kurtosis) were derived from the stop burst using a 1024-point Fast Fourier Transform and a 10ms Hamming window centered on the burst onset. Following [13], prior to the calculation of the spectral moments, all samples were (a) downsampled at 22050 Hz, (b) pre-emphasized with a first order filter with a pre-emphasis coefficient equal to 0.97, and (c) zero-phase high-pass filtered by a 60-tap FIR filter with a cutoff frequency of 70 Hz. The MATLAB code can be found online (http://www.csd.uoc.gr/~kafentz/publication.html).

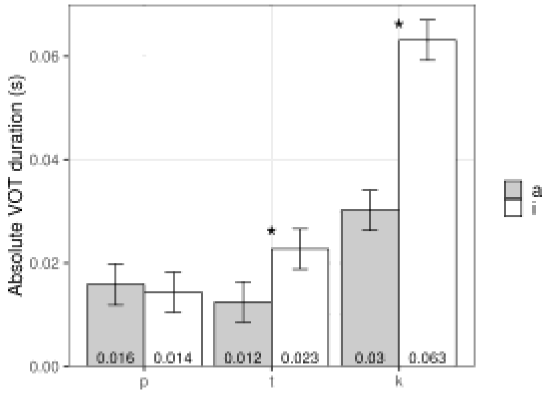
A mixed model ANOVA was carried out for each measure, with place of articulation and vowel as the within-subjects factors, gender and age as the between-subjects factors, and subjects as random factor.

## 3. RESULTS

### 3.1. VOT and closure duration

Significant differences due to place of articulation were found for both absolute ($F(2, 416)=151.32$, $p<0.001$) and relative ($F(2, 416)=392.15$, $p<0.001$) VOT duration. VOT decreased in the order /k/>/t/>/p/. Tukey post hoc tests showed significant differences among all places for the relative measure but not between /p/ and /t/ for absolute duration (abs: /k/=0.046, /t/=0.017, /p/=0.015). In addition, the place of articulation by vowel interaction was significant (abs: ($F(2, 416)=36.93$, $p<0.001$; rel: ($F(2, 416)=77.80$, $p<0.001$); post hoc tests showed that VOT was significantly longer in the /i/ than /a/ context for /k/ and /t/. An increase in duration around 50% is evident in the /i/ context for both /k/ and /t/ in both absolute and relative measures. The

**Figure 1:** Absolute VOT duration in seconds for /p, t, k/ in the /i/ and /a/ vocalic contexts.



palatal [c] was produced with very long VOT (see /k/ in the /i/ context (0.063 sec) in Fig. 1; absolute values are presented in Figure 1 and Table 1 for comparison with the literature). The age main effect was not significant, while gender was significant for the relative VOT measure with males having overall longer VOT than females (F(1, 416)=6.46, p=0.019).

Significant differences due to place of articulation were also found for closure duration (abs: (F(2, 416) = 69.90, p<0.001); rel: (F(2, 416) = 228.47, p<0.001). Closure duration decreased in the order /p/>/t/>/k/. Post hoc tests showed significant differences among all places for the relative measure but not between /p/ and /t/ for absolute duration (abs: /p/=0.117, /t/=0.115, /k/=0.92). The place of articulation by vowel interaction was significant (rel: F(2, 416)=3.59, p=0.028); post hoc tests showed that relative duration was longer in the /i/ context for /t/. A significant age effect showed longer absolute duration for children than adults (F(1, 416)=4.98, p=0.037) but such differences were not found for the relative measure. Gender was not significant. Table 1 presents the absolute duration results.

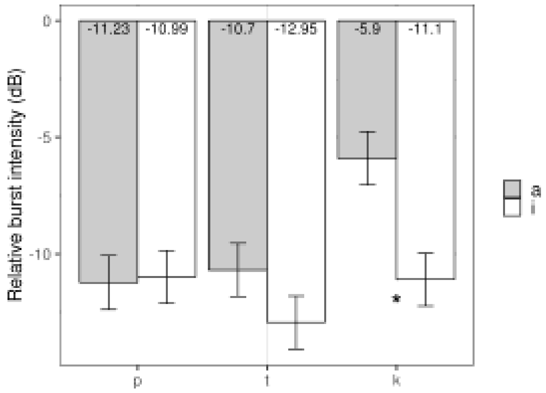
**Table 1:** Mean absolute VOT and closure duration in seconds, and standard deviations in parentheses, for /p, t, k/ for male (M), female (F) adult (A) and children (C) participants.

| | VOT duration | | |
|---|---|---|---|
| | /p/ | /t/ | /k/ |
| MA | 0.014 (0.004) | 0.017 (0.006) | 0.047 (0.020) |
| FA | 0.010 (0.002) | 0.015 (0.007) | 0.043 (0.020) |
| MC | 0.017 (0.008) | 0.022 (0.012) | 0.049 (0.023) |
| FC | 0.021 (0.051) | 0.016 (0.008) | 0.048 (0.025) |
| | Closure duration | | |
| | /p/ | /t/ | /k/ |
| MA | 0.110 (0.022) | 0.106 (0.026) | 0.077 (0.018) |
| FA | 0.114 (0.017) | 0.109 (0.014) | 0.089 (0.020) |
| MC | 0.117 (0.023) | 0.120 (0.031) | 0.101 (0.028) |
| FC | 0.127 (0.029) | 0.128 (0.031) | 0.106 (0.038) |

### 3.2. Relative burst intensity

Place of articulation was significant (F(2, 416)=18.83, p<0.001); post hoc tests showed significantly greater relative burst intensity for /k/ than /p/ and /t/. In addition, significantly greater burst intensity was present in the context of /a/ than /i/ for /k/ (place of articulation by vowel interaction: F(2, 416)=10.79, p<0.001). Fig. 2 shows that the relative burst intensity of /k/ in the /i/ context, i.e. of the palatal [c], is similar to the other stops while in the context of /a/, i.e. velar production, greatest intensity is present. A significant age main effect showed overall greater intensity for adults than children (A=-9.36, C=-11.59), (F(1, 416)=5.93, p=0.024). The gender main effect was not significant.

**Figure 2**: Relative burst intensity in dB for [p, t, k] in the /i/ and /a/ vocalic contexts.

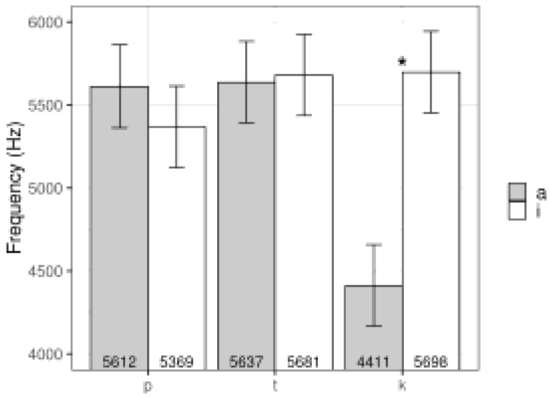



### 3.3. Spectral Moments

#### *3.3.1. Mean frequency*

Place of articulation was significant (F(2, 416)=12.04, p<0.001); post hoc tests showed significantly lower mean frequency for /k/ than /p, t/. The means for the vowel main effect (F(1, 416)=13.35, p<0.001) showed overall lower mean frequency in the /a/ than /i/ context. However, post hoc tests for the place of articulation by vowel interaction (F(2, 416)=19.11, p<0.001) showed significantly lower frequency in the context of /a/ for /k/ only (Fig. 3). No significant age or gender effects were found.

**Figure 3:** Mean frequency in Hz for [p, t, k] in the /i/ and /a/ vocalic contexts.




#### *3.3.2. Standard deviation*

Significant differences in standard deviation (SD) were found as a function of place of articulation (F(2, 416)=34.74, p<0.001); SD decreased in the order /p/>/t/>/k/ (/p/=2278, /t/=2016, /k/=1874). Higher SD was overall present in the context of /a/ than /i/ (F(1, 416)=67.93, p<0.001) with significant differences shown by post hoc tests for /k/ only (/a/=2254, /i/=1495), (place of articulation by vowel interaction (F(2, 416)=29.44, p<0.001). Higher SD was present for female than male speakers (F=2231, M=1882), (F(1, 416)=4.77, p=0.041), for all places of articulation (F(2, 416)=4.17, p=0.016). Higher SD for adults than children was also found (A=2292, C=1820), (F(1, 416)=8.72, p=0.008).

### 3.3.3. Skewness

Place of articulation was significant (F(2, 416)=43.96, p<0.001). Post hoc tests showed significantly higher skewness for /k/ than /p, t/; differences between /p/ and /t/ were not significant (/k/=0.741, /p/= 0.080, /t/= -0.055). The vowel, gender and age effects were not significant.

### 3.3.4. Kurtosis

Place of articulation was significant (F(2, 416)=17.87, p<0.001). Post hoc tests showed significantly higher kurtosis for /k/ than /p, t/; differences between /p/ and /t/ were not significant (/k/=1.47, /t/=0.157, /p/= -0.101). Overall, higher kurtosis was found in the environment of /i/ than /a/ (F(1, 416)=7.03, p=0.008) with significant differences shown by post hoc tests in the environment of /k/ only (/i/=2.25, /a/=0.68), (place of articulation by vowel interaction: F(2, 416)=4.36, p=0.013). The gender main effect was significant (F(1, 416)=6.33, p=0.02); the means showed overall positive kurtosis for the male speakers and negative for the female participants (M=1.02, F= -0.002). No significant age effect was found.

## 4. DISCUSSION

Significant differences in relative closure and relative VOT duration were found among all three stops. Closure duration decreased and VOT duration increased in the order /p/, /t/, /k/ confirming the inverse relationship reported in previous literature on Greek (see [3] for a review) and other languages [7]. Such variation due to place of articulation can be explained on the basis of physiological/ aerodynamic factors as expounded in section 1. It is worth pointing out that significant differences in VOT and closure duration across all stops were evident for the relative measure only, while absolute duration differences between /p/ and /t/ were not significant. This can explain the variability that has been reported for these two Greek consonants in previous work ([2], [14], [18], and [6] for children), mainly on the basis of absolute durations which may be more variable due to rate of articulation changes. In addition, longer absolute and relative VOT duration for /t/ and /k/ and longer relative closure duration for /t/ were found in the context of the high vowel /i/ than the open /a/ (cf. [11], [14], [18]). Such contextual effects on VOT may be interpreted on the basis of aerodynamic factors, i.e. as the supraglottal space is smaller for /i/ than /a/ it takes more time to achieve the pressure differences which are required for the initiation of voicing, hence VOT duration is longer. It is worth pointing out that the longest VOT duration was noted for the palatal [c]. Age differences were not present in VOT duration which is in line with previous developmental work showing that Greek children achieve adult like VOT duration within their second year of life [23], (cf. VOT values for Greek children of similar ages to this study in [6]). An age effect noted for absolute duration with longer closure duration for children than adults was not evident for relative duration, therefore indicating an effect of rate. Finally, longer relative VOT duration was found for male speakers (cf. [25], [27] for English showing longer VOT for females, and [21] for Korean showing longer VOT for males for aspirated stops). Gender related variation may be accounted for on the basis of physiological and sociophonetic factors.

Results on relative burst intensity have shown greater intensity for the velar /k/ in the context of /a/, while for the rest of the consonants, including the palatal [c], intensity was lower (cf. ([10], [28]) for English aspirated stops with greater intensity for /t/ than /k/). An age effect was also found with overall greater burst intensity for adults than children.

The spectral moments analysis showed that the velar /k/ was systematically differentiated from all other places in all four measures. It was characterised by significantly lower mean frequency, lower standard deviation, suggesting a more compact burst spectrum, higher positive skewness indicating negative tilt and thus concentration of energy in the lower frequencies and higher positive kurtosis indicating a more peaked distribution. Negative skewness for /t/ suggests energy concentration in higher frequencies while negative kurtosis for /p/ is indicative of a relatively flat distribution with not so clearly defined peaks. Across the four moments, only standard deviation was significantly different for all different places. For the rest of the measures, /p/, /t/, and commonly [c], did not show significant differences. Finally, there was variation in gender and age effects. Higher standard deviation for female than male speakers and for adults than children suggests overall more diffuse bursts with energy spreading over a larger range of frequencies for female and adult participants. Negative kurtosis for female speakers indicates that overall they have relatively flatter spectra with not so clearly defined peaks.

The work reported in this paper is part of a larger investigation which includes additional parameters, such as analysis of voiced stops, more vowel contexts, vowel transitions, and more subjects, with the last aiming to examine speaker variability in stop production (cf. [8]), and age and gender effects on the basis of a larger sample.

## ACKNOWLEDGMENTS

The data analysed in this paper are from the *SpeakGreek* database, Project title: "SpeakGreek: Developing a biofeedback speech training tool for Greek segmental and suprasegmental features: Application in L2 learning/teaching and clinical intervention", ARISTEIA II, co-financed by ESF and NSFR. Thanks are due to all project team members and all the subjects recorded for the database.


## 5. REFERENCES


[1] Arbisi-Kelm, T., Beckman, M., Kong, E., Edwards, J. 2008. Psychoacoustic measures of stop production in Cantonese, Greek, English, Japanese, and Korean. Poster presented at the *156th Meeting of the Acoustical Society of America*, Miami, Florida.

[2] Arvaniti, A. 2001. Comparing the Phonetics of Single and Geminate Consonants in Cypriot and Standard Greek. *Proc. of the Fourth International Conference on Greek Linguistics*, 37-44.

[3] Arvaniti, A. 2007. Greek Phonetics: The state of the art. *Journal of Greek Linguistics* 8, 97-208.

[4] Botinis, A., Fourakis, M., Prinou, I. 2000. Acoustic structure of the Greek stop consonants. *Glossologia* 11-12, 167-200.

[5] Bundgaard-Nielsen, R.L., Baker, B.J. Kroos, C., Harvey, M., Best, C.T. 2012. Vowel acoustics reliably differentiate three coronal stops of Wubuy across prosodic contexts. *Laboratory Phonology* 3, 133-161.

[6] Chionidou, A., Nicolaidis, K. 2015. Voice onset time production in Albanian and Greek initial voiceless stops. *Albanohellenica* 6, 1-12.

[7] Cho, T, Ladefoged, P. 1999. Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics* 27, 207-229.

[8] Chodroff, E., & Wilson, C. 2017. Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics* 61, 30-47.

[9] Crystal, T., House, A. 1988. Segmental durations in connected speech signals: Current results, *J. Acoust. Soc. Am.* 83, 1553-1573.

[10] Edwards, T. 1981. Multiple feature analysis of intervocalic English plosives. *J. Acoust. Soc. Am.* 69, 535-547.

[11] Efstathopoulou, P. N. 2007. The influence of second language learning on speech production by Greek/English bilinguals. *MA Thesis*. Simon Fraser University. Burnaby, Canada.

[12] Fant, G. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.

[13] Forrest, K., Weismer, G., Milenkovic, P. Dougall, R. N. 1988. Statistical analysis of word-initial voiceless obstruents: preliminary data. *J. Acoust. Soc. Am.* 84, 115-123.

[14] Fourakis, M. 1986. A Timing Model for Word-Initial CV Syllables in Modern Greek. *J. Acoust. Soc. Am.* 79, 1982-1986.

[15] Keating, P. A. 1984. Phonetic and phonological representation of stop consonant voicing. *Language* 60, 286-319.

[16] Kollia, H. B. 2001. Segmental duration changes due to variations in stress, vowel, place of articulation, and voicing of stop consonants in Greek. *J. Acoust. Soc. Am.* 93, 2298.

[17] Nicolaidis, K. 2001 An electropalatographic study of Greek spontaneous speech. *Journal of the International Phonetic Association* 31(1), 67-85.

[18] Nicolaidis, K. 2002. Durational Variability in Vowel-Consonant-Vowel Sequences in Greek: The Influence of Phonetic Identity, Context and Speaker. *Selected Papers on Theoretical and Applied Linguistics from the 14th International Symposium*, Thessaloniki, 280-294.

[19] Nissen, S. L., Fox, R. A. 2009. Acoustic and spectral patterns in young children's stop consonant productions. *J. Acoust. Soc. Am.* 126, 1369-1378.

[20] Nittrouer, S. 1995. Children learn separate aspects of speech production at different rates: Evidence from spectral moments. *J. Acoust. Soc. Am*. 97, 520-530.

[21] Oh, E. 2011. Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics* 39, 59-67.

[22] Ohde, R. N. 1995. Fundamental frequency correlates of stop consonant voicing and vowel quality in the speech of preadolescent children. *J. Acoust. Soc. Am.* 78, 1554-1561.

[23] Okalidou, A., Petinou, K. Theodorou, E. Karasimou, E. 2010. Development of voice onset time in standard-Greek and Cypriot-Greek-speaking preschoolers. *Clinical Linguistics & Phonetics* 24(7), 503-519.

[24] Pickett, J. M., Bunnell, H. T., Revoile, S. G. 1995. Phonetics of intervocalic consonant perception: retrospect and prospect. *Phonetica* 52(1). 1-40.

[25] Ryalls, J., Zipprer, A., Baldauff, P. 1997.A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research* 40, 642-645.

[26] Stevens, K. N., Blumstein, S. E. 1978. Invariant cues for place of articulation in stop consonants. *J. Acoust. Soc. Am.* 64, 1358-1368.

[27] Whiteside, S. P., Irving, C. J. 1998. Speakers' sex differences in voice onset time: A study of isolated word production. *Perceptual and Motor Skills* 86, 651-654.

[28] Zue, V. W. 1976. Acoustic characteristics of stop consonants: a controlled study. *MIT Technical Report* 523, Lincoln Laboratories.