VOICING CONTRAST IN TURKISH: SIMULTANEOUS MEASUREMENTS OF ACOUSTICS, EPG AND INTRAORAL PRESSURE

Özlem Ünal-Logacev[a,b] and Susanne Fuchs[b]
[a]Anadolu University; [b]Zentrum für Allgemeine Sprachwissenschaft, Berlin Germany
ozlemonal@anadolu.edu.tr; fuchs@zas.gwz-berlin.de

ABSTRACT

The aim of this study was to investigate the Turkish voicing contrast with simultaneous measurements of acoustics, EPG and intraoral pressure. We further aimed to understand the mechanisms behind the maintenance or dis-appearance of voicing in Turkish. According to our results, intraoral pressure rises more slowly for voiced sounds than for their voiceless counterparts. Voiced fricatives have shorter duration, more palatal contact, and lower intraoral pressure velocity maximum compared to voiceless cognates. Furthermore, we found a positive correlation between intraoral pressure and the number of (relative) anterior palatal contacts.

Keywords: Turkish, voicing contrast, EPG, intraoral pressure.

1. INTRODUCTION

A challenge in the investigation of the phonological voicing contrast is that the maintenance or disappearance of voicing can be achieved by various articulatory maneuvers. This phenomenon has been called *motor equivalence* and can roughly be defined as the capacity to perform the same task using different strategies.

For example, several articulatory maneuvers have been described which can serve the maintenance of voicing during a vocal tract closure or constriction: laryngeal lowering [1], velar elevation, tongue root advancement [2], and jaw lowering [3].

From a technical point of view, the use of invasive techniques is required to record the entire complex of maneuvers. However, one possible way to avoid their usage is to measure the intraoral pressure (PIO) during articulation since it indirectly reflects the cumulative effect of supralaryngeal movements on the maintenance or disappearance of voicing. PIO measurement can further be combined with articulatory measures of tongue-palatal contact patterns using *Electro-palatography (EPG)* [4]. In the following, we will present an investigation into the voicing contrast in Turkish using both measurement techniques. The Turkish language is of particular interest because the voicing properties of Turkish are not sufficiently explored. For example, there is disagreement on whether Turkish has a three-way or a two-way voicing contrast [5, 6, 7, 8]. Several phonetic properties of voiced and voiceless sounds may lead to such disagreements.

Moreover, we are not aware of any objective articulatory or aerodynamic measurements of the Turkish voicing contrast so far. Such data could help improve therapeutic interventions, in particular in cleft palate speakers who have severe problems in producing voiced and voiceless obstruents. In such therapies, verbal descriptions of articulatory strategies, and tactile feedback cues are provided, and would be more efficient if informed by articulatory data.

In our investigation, we wanted to answer the following questions:

i) How do speakers of Turkish realize the voicing contrast in obstruents?
ii) What mechanism underlies the maintenance or disappearance of voicing? More specifically, what is the relation between acoustic voicing in combination with intraoral pressure rise and tongue palatal contact patterns?

We expected the intraoral pressure slopes to rise more slowly in phonologically voiced obstruents than in voiceless cognates, as has been found Zygis et al. [9] for Polish and German. The simultaneous measurements performed in this study will additionally allow us to investigate the relationship between intraoral pressure and articulatory measures, which have not been provided in [9]. We expected a close correlation between intraoral pressure rise and the amount of tongue palatal contact patterns, as reported in [4] for German voiceless obstruents and will add the phonologically voiced counterparts.

2. METHODS

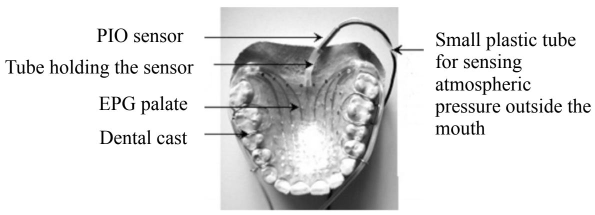
2.1. Participants

The experiment was conducted with six Turkish native speakers (age range: 25-38; 3 females and 3 males) without any known disorders of speech, language or hearing.

2.2. Experimental set-up

Data was recorded with three different systems simultaneously: i) the acoustic signals were recorded on DAT at a sampling rate of 48 kHz, ii) the EPG data was recorded by a Reading EPG 3 system at a sampling rate of 100 Hz, iii) the intraoral pressure signal was recorded with a pressure sensor (Endevco 8507C-2) attached to the posterior end of the EPG palate (cf. Figure 1). The sensor measured the difference between atmospheric pressure and intraoral pressure via a small tube passing through the teeth outside the oral cavity. The PIO signal was sampled to 6000 Hz.

Figure 1: Intraoral pressure sensor attached to the posterior end of an EPG palate.



2.3. Speech stimuli and procedure

This study was conducted as part of a larger experiment that investigated speech production in Turkish. Over the course of the experiment, participants read five randomized lists with 53 sentences. That is, each sentence was read five times in different positions in the list. 16 sentences in each list were part of the present experiment, while 37 sentences belonged to other experiments.

All sentences consisted of the same carrier phrase combined with different two-syllable target words, as illustrated in Example 1. The target word was placed in the second position to avoid prosodic influences.

(1) Arda *tava* anlamlı bir sözcüktür dedi.
(Arda said (that) *'pan'* is a meaningful word.)

Each target word for the present study contained one of eight target sounds (/t, d, s, z, ʃ, ʒ, tʃ, dʒ/). Each target sound was followed by two different vowels (/a/ and /i/) in different words, thus yielding a total of 16 target words.

Participants were instructed to read each sentence aloud at a normal speech rate. Every participant wore the artificial palates with attached PIO pressure sensor at least for 30 minutes before the experiment started.

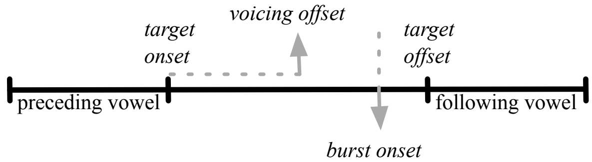
2.4. Data analyses

Overall, we analyzed 480 tokens (80 per participant). Three data points from the EPG data, and 14 data points from the PIO data were excluded due to misreading, recording or annotation problems.

2.4.1. Acoustic annotations

We first manually annotated onsets and offsets of target consonants, as well as the previous and following vowels. This was done in Praat (version 5.3.53) [10]. We also labeled voicing offsets for all target sounds and the time of burst for plosives and affricates as shown in Figure 2.

Figure 2: Annotation for stops and affricates.



We calculated the voicing duration, closure duration, percentage of voicing and fricative duration from the annotated time points as follows:

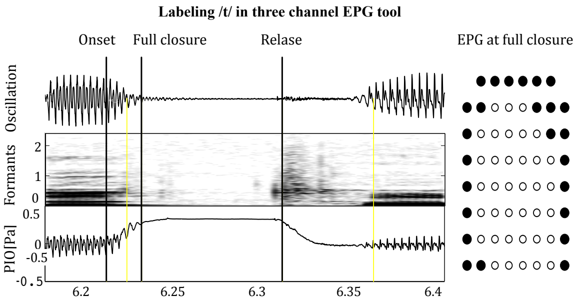
(1) Voicing Duration=*Voicing Offset–Target Onset*
(2) Closure Duration=*Burst–Target Onset*
(3) Voicing Percentage $= \frac{Voicing\ Duration}{Closure\ Duration} \cdot 100$
(4) Fricative Duration= *Target Offset–Target Onset*
(5) Voicing Percentage $= \frac{Voicing\ Duration}{Fricative\ Duration} \cdot 100$

When the voicing continues over the course of the production of a sound, the percentage of voicing is considered to be 100%. When there is no voicing during the production of a sound it is considered to be 0%. For some cases a burst did not exist and continuous voicing was visible throughout the closure. In these cases we used the following vowel's onset to calculate closure duration.

2.4.2. EPG annotations

After labeling the recordings in Praat, we imported the acoustic data with annotations into a MATLAB based tool written by Mark Tiede in order to label the EPG data. For each trial, we determined the *onset of the target sound* in EPG as the earliest time point after the previous vowel at which two additional electrodes were activated. Figure 3 shows an example annotation of the EPG data with the cursor at the full closure point.

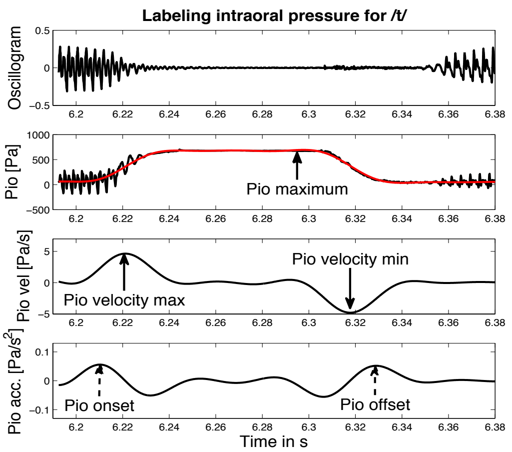
Figure 3: EPG tool with acoustics, spectrogram, intraoral pressure data from top to bottom and EPG (right). The vertical dark lines correspond to annotations of onset, maximum and release of a stop.



We annotated the time of *full closure* and *release* for plosives and affricates (/t, d, tʃ, dʒ/), as well as the time when the maximum number of electrodes was activated for fricatives. Release was determined as the deactivation of an electrode following a full closure. After labeling, the following EPG indices were calculated for the time points of interest (onset, full closure, release): PC (percent of contact), ANT (contact in the anterior region), COG (center of gravity) [18].

2.4.3. *PIO annotations*

Figure 4. Annotation of PIO onset, PIO offset, maximum pressure, and velocity extreme for /t/.



Intraoral pressure data were analyzed using MATLAB (7.13). The raw intraoral pressure data were first filtered and then the first and second derivatives were calculated. Based on these two, we defined different temporal landmarks, which are shown in Figure 4.

2.5. Statistical analysis

To test for effects of voicing on target duration, PC, ANT, PIO maximum, PIO velocity maximum, we used linear mixed-effects models [11, 12, 13] using the lme4 package [14] in R [15]. We included the voicing duration in percentage and manner of articulation as fixed effects. We standardized our predictor variable by participants (centered and divided by one standard deviation). We added random intercepts by participant and list randomization, as well as by-participant random slopes for voicing duration. Categorical predictors were coded with treatment contrasts.

3. RESULTS

Table 1 shows the means and standard deviations of the acoustic measures obtained in the experiment (percentage of voicing, target duration), as well as EPG (PC, ANT and COG) and PIO variables (PIO max and velocity maximum).

Table 1: Mean and SD values in brackets; manner categories with voiced (v) and voiceless (vl) phonemes in columns and parameters in rows.

	Plosives		Fricatives		Affricates	
	vl	v	vl	v	vl	v
Acoustics						
Voicing	19	97	12	91	25	94
%	(13)	(10)	(5)	(20)	(17)	(16)
Duration	145	98	137	108	160	123
	(52)	(39)	(27)	(26)	(38)	(34)
EPG						
PC	42	39	41	46	60	60
	(6)	(7)	(7)	(7)	(9)	(9)
ANT	58	54	47	57	79	81
	(10)	(11)	(13)	(10)	(11)	(9)
COG	4.9	4.9	4.2	4.4	4.6	4.6
	(0.2)	(0.3)	(0.6)	(0.5)	(0.2)	(0.2)
PIO						
Piomax	706	413	539	424	709	605
	(161)	(173)	(141)	(152)	(108)	(168)
Velocity	3.9	1.6	2.4	1.2	3.9	1.6
	(1.1)	(0.5)	(0.8)	(0.4)	(1.0)	(0.4)

According to Table 1, voiceless sounds have a longer overall duration than voiced sounds, while affricates have longer durations than fricatives and plosives. Affricates have more ANT and PC contact compared to the other sounds. The voiceless plosive /t/ and the affricate /tʃ/ have the highest pressure and PIO velocity maximum values. PIO velocity is also higher for all voiceless sounds compared to their voiced counterparts.

We found a strong negative correlation between voicing percentage and PIO velocity maximum (R= -0.69, CI=[-0.73, -0.64]) such that the longer the voicing duration, the lower the PIO velocity. Furthermore, voicing percentage was negatively correlated with target duration (R=-0.57, CI=[-0.63, -0.51]) and with PIO maximum (R=-0.49, CI=[-0.56, -0.42]). This means that voiced sounds had shorter durations and lower intraoral pressure values.

Furthermore, there was a positive correlation between PIO maximum and COG (R=0.35, CI=[0.26, 0.42]), and between PIO maximum and ANT (R=0.40, CI=[0.32, 0.47]) values.

We used linear mixed-effects models to test for effects of voicing percentage on target duration, PC, ANT, PIO maximum and PIO velocity maximum (Table 2).

Table 2: Summary statistics of several linear mixed effects models. We omitted the main effect of manner of articulation.

		Est. (β)	SE	t
Duration	Intercept	142.98	12.03	11.88
	Voicing	-16.90	2.46	-6.88
	Voicing×Pl−Aff	-7.12	3.13	-2.28
	Voicing×Fr−Aff	3.20	2.72	1.18
PC	Intercept	59.92	1.81	33.06
	Voicing	0.32	0.64	0.50
	Voicing×Pl−Aff	-1.76	0.85	-2.07
	Voicing×Fr−Aff	1.82	0.74	2.45
ANT	Intercept	79.84	1.82	43.92
	Voicing	0.99	1.04	0.95
	Voicing×Pl−Aff	-2.89	1.42	-2.04
	Voicing×Fr−Aff	3.21	1.23	2.60
PIOmax	Intercept	660.54	41.86	15.78
	Voicing	-56.59	19.13	-2.96
	Voicing×Pl−Aff	-88.41	15.78	-5.60
	Voicing×Fr−Aff	-10.49	13.74	-0.76
Velocity	Intercept	2.89	0.16	17.66
	Voicing	-1.20	0.13	-8.88
	Voicing×Pl−Aff	0.03	0.08	0.38
	Voicing×Fr−Aff	0.60	0.07	8.45

The linear model showed that voiced sounds have shorter durations than voiceless sounds (β=−16.90). We also found that voiced plosives have shorter durations than affricates (β=−-7.12).

We did not find a main effect of voicing on PC or ANT. However plosives in comparison to affricates have fewer PC contacts (β=−-1.76) and fewer ANT contacts (β=−2.89) when the voicing increases. The picture is reversed for fricatives: fricatives have more PC and ANT contacts compared to the affricates when the voicing increases (β=1.82, and β=3.21).

Voicing duration has a negative effect on the PIO maximum (β=−56.59) and voiced plosives have a lower PIO maximum compared to voiced affricates (β=−88.41). Among all variables, voicing has the largest effect on the PIO velocity maximum: with increased voicing PIO velocity maximum values decreases (β=−1.20). Moreover, PIO velocity increases for fricatives (β=0.60) compared to affricates when the voicing increases.

4. DISCUSSION

In this paper, we aimed to investigate the Turkish voicing contrast in obstruent sounds. We found longer durations for voiceless obstruents compared to voiced cognates, as reported in the Kopkallı-Yavuz study [6]. We also found a difference in the number of palatal contacts between voiced and voiceless sounds for fricatives. This finding is consistent with previous research on English. For instance, McLeod [16] found a significantly greater amount of tongue-palate contact for /z/ compared to /s/ in the word initial position. Yoshioka [17] found very similar differences between the same sounds in word medial position in whispered speech. In another study [4], more anterior contacts were found in phonologically voiced alveolar fricatives than in voiceless cognates. Thus, our findings are in line with previous findings regarding fricatives.

Plosives are showing the opposite pattern to fricatives. The amount of palatal contact *decreases* for plosives in comparison to affricates when the voicing duration increases.

We further found that voiced sounds have lower intraoral pressure maxima and that this effect is more pronounced for plosives.

Moreover, our results concerning the relationship between voicing and the PIO velocity maximum is in line with Zygis et al.'s [9] findings that intraoral pressure slopes rise more slowly for voiced sounds.

Our second aim was to determine the relationship between the intraoral pressure rise and the type of tongue-palate contacts. We found that the PIO maximum increases with more anterior contacts (COG and ANT have positive correlation with the PIO maximum), which is consistent with the findings in [4] for German voiceless obstruents.

5. REFERENCES

[1] Ewan, W. G. and Krones, R. 1974. Measuring larynx movement using the thyroumbrometer. *J. of Phon.* 2, 327-335.

[2] Westbury, J. R. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *J. Acoust. Soc. Am*. 73, 1322–1336.

[3] Fujimura, O. and Miller J.E. 1979. Mandible height and syllable-final tenseness. *Phonetica* 36, 263-272.

[4] Fuchs, S., and Koenig, L.L. 2009. Simultaneous measures of electropalatography and intraoral pressure in selected voiceless lingual consonants and consonant sequences of German. *J. Acoust. Soc. Am*. 126, 1988-2001.

[5] Öğüt F., Kılıç M. A., Engin E. Z., Midilli R. 2006. Voice onset times for Turkish stop consonants. *Speech Communication* 48, 1094-1099.

[6] Kopkallı-Yavuz, H. 2000. Studies on Turkish and Turkic languages. In: Göksel, A., Kerslake, C. (eds), *Proceedings of the Ninth International Conference on Turkish Linguistics*, Lincoln College, Oxford, Volume 46 of Turcologica Series.

[7] Wilson, S. 2003. A phonetic study of voiced, voiceless and alternating stops in Turkish. *CRL Newsletter* 15, 3–13.

[8] Kallestinova, E. 2004. Voice and aspiration of stops in Turkish. *Folia Linguistica* 38, Issue 1-2, 117–144.

[9] Zygis, M., Fuchs, S., Koenig, L. 2012. Phonetic explanations for the infrequency of voiced sibilant affricates across languages. *J. of Lab. Phon.* 3, 299-336.

[10] Boersma, P. & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program]. Version 5.3.53 retrieved 04 October 2013 from http://www.praat.org/).

[11] Pinheiro, J.C., and Bates, D.M. 2000. *Mixed-Effects Models in S and S-PLUS*. New York: Springer.

[12] Gelman, A., Hill, J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.

[13] Baayen, R. H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R.* Cambridge: Cambridge University Press.

[14] Bates, D., Maechler, M., Bolker, B., Walker, S. 2013. *lme4: Linear Mixed-Effects Models Using Eigen and S4.* R package.

[15] R Development Core Team. 2013. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

[16] McLeod, S., Roberts, A., Sita, J. 2006. Tongue/palate contact for the production of /s/ and /z/. *Clinical Linguistics and Phonetics* 20(1), 51–66.

[17] Yoshioka, H. 2008. The role of tongue articulation for /s/ and /z/ production in whispered speech. *Acoustics '08 Paris Proceedings,* 2335-2338. Paris: Societé Française d'Acoustique.

[18] Hardcastle, W.J., Gibbon, F. and Nicolaidis, K. 1991. EPG data reduction and their implications for studies of lingual coarticulation. *Journal of Phonetics.* 19, 251-266.