Eastern Armenian is an Indo-European language with a three-way laryngeal contrast for stops: voiced, voiceless unaspirated, and voiceless aspirated, e.g., /b, p, pʰ/. Word-final voiceless unaspirated stops are variably ejectives: /p, pʰ/. To analyze the phonetics of final ejectives, we recorded three speakers reading from a word list of 51 words. The words had the three types of stops, at three places of articulation, at different syllable positions (initial, final, intervocalic). We measured the target stops in terms of their acoustics (VOT) and aerodynamics (oral airflow, intraoral pressure, and volume of exhaled air). We likewise recorded EGG signals for their articulation. We found that all three speakers almost always used ejectives for word-final voiceless unaspirated stops. We suggest our speakers use ejectives as a fortition process that creates a clearer phonetic contrast between final (ejectives) voiceless unaspirated stops vs. final voiceless unaspirated stops.

**Keywords:** Armenian, aspiration, ejectives, aerodynamics, EGG

### 1. INTRODUCTION

Eastern Armenian is an Indo-European language. It forms a separate branch in the IE family, with two standardized varieties (Western and Eastern) that differ in the laryngeal contrasts of stops and affricates [2, 3, 4, 5, 6, 7]. Eastern has a three-way contrast between voiced, voiceless unaspirated, and voiceless aspirated stops [8, 9]: bilabial /b, p, pʰ/, coronal /d, t, tʰ/, and velar /ɡ, k, kʰ/. The three-way contrast has been confirmed in previous acoustic studies in terms of voice onset time (VOT) [10, 11, 12, 13, 14, 15]. Voiced stops have negative VOT, voiceless unaspirated have zero VOT or a short lag, and voiceless aspirated stops have long positive VOT. This paper replicates this finding.

### 2. BACKGROUND

Outside of Armenian, there is extensive work on the phonetics of ejectives [16, 17, 18, 19, 20, 21, 22]. For Eastern Armenian, some phonological and phonetic studies report that some speakers variably use ejectives for word-final voiceless unaspirated stops [23, 24, 25]. For example, the word կատակ ‘joke’ is typically pronounced with a final voiceless unaspirated stop [k]: kɑtɑk. However, some speakers produce an ejective instead of a final voiceless pulmonic stop: katak'. It is unclear how widespread the use of final ejectives is in Eastern Armenian. Some studies had participants who consistently used ejectives [26], never used ejectives [13], or a subset variably did [1, page 67]. One study [14] reports that 2 out of 8 speakers of Eastern Armenian (from Yerevan, Armenia) produced final ejectives. Another study [27] reports that 4 out 225 tokens of voiceless unaspirates were ejectives in their sample. The lack of final ejectives is reported to be the prescriptive norm [28, page 17].
3. METHODOLOGY

Three speakers were recorded reading from a word list. All three were native speakers of Eastern Armenian. Their age range was 25-38. Two speakers (M1 and M2) were male, born and raised in Armenia. One speaker (F3) was female, born and raised in Iran. The two regional dialects are not known to have significant differences for the laryngeal contrast [13, 29].

The stimuli was a word list of 51 words. Words were uttered in isolation without any carrier phrase or sentence. For each word, we asked the speaker to repeat the word three times. Some speakers would sometimes repeat 4 or 6 times by accident. Each word had one or more of the following target stops: /p, t, k, b, d, g/. The stops were in different syllable positions. We report on a subset of these 51 words where where the target stops were found in one of the following positions: word-initial pre-vocalic (#_V), intervocalic (V_V), or word-final post-vocalic (V_#). In Table 2, we provide the recorded tokens per speaker and per context. Some stop-context combinations were sparsely populated, such as non-initial voiced stops.

Table 2: Tokens per speaker, context, and stop

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Coronal</th>
<th>Velar</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b p pʰ</td>
<td>d t tʰ</td>
<td>g k kʰ</td>
<td></td>
</tr>
<tr>
<td>#_V M1</td>
<td>9 6 6 4</td>
<td>9 3 9 4</td>
<td>12 3 6 4</td>
<td>65</td>
</tr>
<tr>
<td>M2</td>
<td>9 6 6 4</td>
<td>9 3 9 4</td>
<td>12 3 6 4</td>
<td>63</td>
</tr>
<tr>
<td>F3</td>
<td>9 12 6 4</td>
<td>9 3 9 4</td>
<td>12 3 6 4</td>
<td>69</td>
</tr>
<tr>
<td>Total</td>
<td>27 24 18</td>
<td>27 9 27</td>
<td>36 9 18</td>
<td>195</td>
</tr>
<tr>
<td>V_V M1</td>
<td>9 6 4</td>
<td>3 3 6 4</td>
<td>3 3 3 4</td>
<td>36</td>
</tr>
<tr>
<td>M2</td>
<td>6 6 6 4</td>
<td>3 3 3 4</td>
<td>3 3 3 4</td>
<td>36</td>
</tr>
<tr>
<td>F3</td>
<td>12 6 4</td>
<td>3 3 6</td>
<td>4 3 4</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>27 18</td>
<td>9 9 18</td>
<td>10 9</td>
<td>109</td>
</tr>
<tr>
<td>V_# M1</td>
<td>3 3</td>
<td>6 3</td>
<td>12 3</td>
<td>30</td>
</tr>
<tr>
<td>M2</td>
<td>3 3</td>
<td>6 3</td>
<td>12 3</td>
<td>30</td>
</tr>
<tr>
<td>F3</td>
<td>6 3</td>
<td>6 3</td>
<td>12 3</td>
<td>33</td>
</tr>
<tr>
<td>Total</td>
<td>12 9</td>
<td>18 9</td>
<td>36 9</td>
<td>93</td>
</tr>
<tr>
<td>Total</td>
<td>27 63 45</td>
<td>36 36 34</td>
<td>45 55 36</td>
<td>397</td>
</tr>
</tbody>
</table>

Participants were recorded by the first author in Paris 2017. Speaker F3 was recorded in a soundproof room of the Institut de Linguistique et Phonétique Générales et Appliquées at the Sorbonne Nouvelle university. Speakers M1-M2 were recorded in a quiet school-room. We recorded acoustic (VOT), aerodynamic (oral airflow, intraoral pressure, volume of exhaled air), and articulatory data. To get articulatory data, we used electroglottography (EGG) to visualize the contacts of the vocal folds [30]. To get aerodynamic data, we used a portable machine EVA2.

Our recording protocol was as follows. We placed an EGG device on the neck of the speaker. We asked the speaker to insert a 2mm diameter plastic tube inside the mouth. We attached a flexible rubber mask on the speaker’s mouth. To avoid any air leakage, we used two types of masks, one for the female speaker and one for the male speakers. The speaker was asked to avoid blocking the tip of the tube as much as possible. We attached an inlet of the plastic tube to an EVA2. We ran a trial on the speaker to make sure that the equipment was stable. We recorded the stimuli where the target sound is a bilabial. We removed the tube that was in the mouth. We continued with the rest of the stimuli.

The signals were analyzed using the Phonedit Signiax software which is developed by the Aix-en-Provence Speech and Language Laboratory [31]. We used this software to record, visualize, and display multi-parameter data. Acoustic analysis was done with Praat [32]. Word-final segments were measured by taking the period between the first and last cycles of the signal.

4. RESULTS

All our word-final V_# tokens of /p, t, k/ were ejectives. This is likely because the stimuli were said in isolation, encouraging hyper-articulation. When discussing our measurements, we focus on the bilabial series for illustration.

4.1. Acoustic measurements: VOT

VOT was measured for the stops in all three syllable positions. Results are in Table 3, averaged across speakers. For measuring word-final VOT, we measured between the first and last cycles of the signal.

Table 3: Voice onset time per context and stop (in milliseconds)

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Coronal</th>
<th>Velar</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b p pʰ</td>
<td>d t tʰ</td>
<td>g k kʰ</td>
<td></td>
</tr>
<tr>
<td>V_# M1</td>
<td>12.43 16.59 20.59</td>
<td>55.68 49.65 50.10</td>
<td>19.97 45.79</td>
<td>9.52</td>
</tr>
<tr>
<td>M2</td>
<td>12.43 16.59 20.59</td>
<td>55.68 49.65 50.10</td>
<td>19.97 45.79</td>
<td>9.52</td>
</tr>
<tr>
<td>F3</td>
<td>12.43 16.59 20.59</td>
<td>55.68 49.65 50.10</td>
<td>19.97 45.79</td>
<td>9.52</td>
</tr>
<tr>
<td>Total</td>
<td>12.43 16.59 20.59</td>
<td>55.68 49.65 50.10</td>
<td>19.97 45.79</td>
<td>9.52</td>
</tr>
</tbody>
</table>

In all three syllable positions (initial, intervocalic, and final), the stops are distinguished in terms of negative, zero, and positive VOT. This is expected. Place has an effect on VOT values [10, 33, 19].

In the word-final position, we find clear VOT differences across 2 of the 3 laryngeal categories. Our stimuli lacked cases of word-final voiced stops. We impressionistically perceived that all three speakers pronounced word-final voiceless unaspirates as ejectives or as ejectives.
4.2. Aerodynamic measurements

4.2.1. Intraoral pressure

Intraoral pressure was measured in hPa for the bilabial series. Results are in Table 4.

Table 4: Intraoral pressure for bilabials per context (in hPa)

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>p</th>
<th>pʰ</th>
</tr>
</thead>
<tbody>
<tr>
<td>#_V</td>
<td>Avg (Std)</td>
<td>4.52 (1.31)</td>
<td>5.07 (2.1)</td>
</tr>
<tr>
<td>V_#</td>
<td>Avg (Std)</td>
<td>7.13 (1.73)</td>
<td>7.58 (1.32)</td>
</tr>
<tr>
<td>V_V</td>
<td>Avg (Std)</td>
<td>7.72 (3.77)</td>
<td>6.83 (1.66)</td>
</tr>
</tbody>
</table>

Within the bilabial series, first consider word-initial stops. the voiced stop /b/ has the lowest intraoral pressure. The pressure is increased for the other two voiceless stops. For voiceless /p/, we see a rapid rise in pressure during closure (around 5 hPa), and then the pressure drops after the closure ends [34, 35]. For /pʰ/, we also find that the pressure increases during the closure, and then reaches its maximum during the release (around 6 hPa) [16]. Pressure is then dropped to around 1.4 dm³/s.

Next consider intervocalic /pʰ/. During closure, the intraoral pressure rises until release. At the release, the vocal folds are in contact, making the intraoral pressure have a zigzag shape, and making the EGG signal be not flat.

Finally consider final /p/ which surfaces as ejectives [pʰ] as in [kʰɑɾpʰ] ‘swan’. After the vowel, the vocal folds spread to let the air pass. This causes an increase in intraoral pressure. Then the vocal folds close. During closure, the intraoral pressure increases until it reaches its maximum. At release, vocal fold vibration is stabilized.

4.2.2. Oral airflow

Oral airflow was measured in dm³/s for all stops. Results are in Table 5.

Table 5: Oral airflow per context and stop (in dm³/s)

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Coronal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>p</td>
<td>pʰ</td>
</tr>
<tr>
<td>#_V</td>
<td>0.33</td>
<td>0.34</td>
<td>0.31</td>
</tr>
<tr>
<td>V_#</td>
<td>0.28</td>
<td>0.68</td>
<td>0.18</td>
</tr>
<tr>
<td>V_V</td>
<td>0.38</td>
<td>0.91</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Consider word-initial bilabials. we see a rapid rise in airflow for /p/ at 0.56 dm³/s. We find that the rise is strongest for the aspirated stops like /pʰ/; this is expected [36].

For intervocalic /pʰ/ the air flow reaches approximately 0.60 dm³/s.

For the aspirate [pʰ], the release is early, gradual, with significant airflow. But for the ejective [pʰ], the release is later; which implies that the delay of the release is due to the compress of the air between the closed larynx and the lips. Hence, when the air is released, the releasing is more sudden, with nearly no airflow.

Word-finally, the voiceless unaspirated stops are ejective. They differ from the aspirated stops because the latter have a higher air flow.

Across the three places, the coronal stop [tʰ] displays some differences. During the articulation of final [tʰ], we see that the glottis is still adducted and that the airflow is relatively low. The noise during release is smaller than that of the bilabial [pʰ] and velar [kʰ]. For velar [kʰ], the noise upon release is more abrupt and sounds more intense.

4.2.3. Air volume

During the VOT, we calculated the air volume by taking the integral of the oral air flow. We treat the air volume as the amount of pulmonary air expelled from the mouth, per unit of time measured in dm³/s (Table 6).

Table 6: Air volume per context and stop (in dm³/s)

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Coronal</th>
<th>Velar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>p</td>
<td>pʰ</td>
</tr>
<tr>
<td>#_V</td>
<td>0.0021</td>
<td>0.0151</td>
<td>0.0022</td>
</tr>
<tr>
<td>V_#</td>
<td>0.0026</td>
<td>0.0157</td>
<td>0.0023</td>
</tr>
<tr>
<td>V_V</td>
<td>0.0024</td>
<td>0.0157</td>
<td>0.0023</td>
</tr>
</tbody>
</table>

Consider word-initial bilabials. the largest amount of air is expelled for the voiceless aspirated stop /pʰ/ than for its voiced and unaspirated counterparts /b, p/. Thus, aspirated stops have the highest air volume.

Next consider intervocalic /pʰ/. Air volume is directly proportional to air pressure for 2/3 of the speakers. During VOT, the pressure reaches its maximum.

Word-finally, the voiceless unaspirated stops are ejective. They differ from the aspirated stops because the latter have a higher air volume.

4.3. Articulatory properties: EGG

We recorded the EGG signal as well. We provide illustrations of the EGG recordings in Figure 1 for the word [ɪsʊɑɾpʰ] with [pʰ] and [kʰɑɾpʰ] with an ejective [pʰ], produced by speaker M2. We provide the charts for the intraoral pressure and oral airflow. The audio is peak clipped, but that should not affect
measurement of VOT.

**Figure 1**: Contrasting final aspirated [pʰ] vs. ejective [p’]

[tsovɑpʰ] ‘seashore’ ęmļuqəł [kɑɾɑp’] ‘swan’ ɥu̯pʊnʊŋ

Waveform

Intraoral pressure (hPa)

Oral airflow (dm³/s)

EGG

The first screenshot of the EGG for the aspirated final [pʰ] implies simultaneous opening of the glottis and the release of the air. However, at the second, for the ejectivized final [p’], the glottis is held closed for a while, the pressure is raised during the closure of the glottis, and then there is a sudden release.

**5. DISCUSSION**

Our speakers produced final voiceless aspirated stops as ejectives. In contrast, for prescriptive Eastern Armenian speech, the norm is to lack ejectivization [28, page 17]. Acoustically, the final voiceless unaspirated stops have short-lag VOT. Thus they are unaspirated acoustically. But based on the aerodynamic data, they act as ejectives.

In our phonological interpretation, final ejectivization seems to be too speaker-variable to be allophonic. However, final ejectivization is not phonologically contrastive. It seems that final ejectivization is an optional fortition process in Eastern Armenian. Phonetically, we suggest that ejectivization acts as a way to phonetically enhance the contrast between final voiceless unaspirated stops and final voiceless aspirated stops. Glottal reinforcement of final voiceless stops is cross-linguistically attested [37]. Through ejectivization, the voiceless status of the final voiceless unaspirated stops is enhanced by the anticipation of glottal closure. During the stop closure, the glottal pressure occurs. The ejectivization does not replace the sounds [p, t, k]; it reinforces them.

This interpretation has been suggested in the past for Armenian. [1, page 67] demonstrate that the difference between voiceless unaspirated and aspirated stops is due to the force of the release. The unaspirated stops are released more “weakly”, or sometimes they are not released at all. In contrast, the aspirates have a relatively shorter closure duration, and then release, then noise and a flow of oral air. For voiceless unaspirated stops, they say that Armenian speakers “probably” have glottal closure during articulation and therefore sometimes slightly ejectivize.

**6. CONCLUSIONS**

Eastern Armenian has a 3-way laryngeal contrast between voiced, voiceless unaspirated, and voiceless aspirated stops. The language is reported to have word-final ejectivization of voiceless unaspirated stops, as a fortition process. Such a process is speaker-specific and highly variable across speakers and geographic regions.

Based on our data, we found three speakers who ejectivize their final voiceless unaspirated. Glottal closure exists during the articulation of these final voiceless unaspirates. The ejectivized stops have noise during release, and the amount of air flow is not great. From an articulatory point of view, the glottis is closed during the production of these stops, thus triggering ejectivization.

From an articulatory point of view, the ejectivized consonants are extensions of the voiceless unaspirated consonants. And simultaneously from an acoustic angle, the ejectivized consonants [p’, t’, k’] are in contrast with their aspirated counterparts [pʰ, tʰ, kʰ] in the word-final position.

**7. ACKNOWLEDGEMENTS**

We thank Dr. Didier Demolin, Sipana Tchakerian, M. Rosario Signorello, and M. Nail Aras for help in setting up the experiment in Paris. We thank Scott Seyfarth and Canaan Breiss for helping us with guiding us though the literature, and we thank Jeremy Steffman for help in interpreting the data.

**8. REFERENCES**


1 The coronal series is typically dental, but some speakers use alveolar stops.
2 Victoria Khurshudyan informs us that word-initial ejectivization is likewise attested in natural speech. Though we’ve been unable to find such cases in our data.
3 We also recorded a fourth speaker. We had to set aside their recordings because we later learned they had ear surgery and their voice was altered.
4 http://www2.lpl-aix.fr/lpldev/phonedit/