Articulatory Constraints and Coronal Stops: an EPG study

Mitsuhiro Nakamura
Nihon University, Tokyo, Japan
m-nakamura@eco.nihon-u.ac.jp

Abstract

This study explores articulatory and coarticulatory properties of the coronal stop articulations [t, d, r] in Japanese, using the technique of electropalatography (EPG). We examine the spatiotemporal characteristics of linguopalatal contact patterns, closure duration, and intergestural coordination, relating them to factors of the articulatory control mechanisms. The results bear on issues in models of lingual articulation, the concept of articulatory constraints and phonological feature specifications.

1. Introduction

This study explores articulatory and coarticulatory properties of the coronal stop articulations [t, d, r], which are practically described by the presence, or absence, of voicing and the aspectual process ‘tapped’ [1]. Particularly, we attempt to uncover where the distinctive properties of [r] originate from and how they are related to those of [t, d].

Previous studies have shown that the closure duration and the amount of tongue contact are controlled consonant-specifically for [t, d] and for [d, r][2, 3], but few attempts have been made for the relation between the three. It is reasonable to expect that the three stops constitute a hierarchy in terms of the contact amount and the closure duration: [t]>[d]>[r]. The question arises as to the significance of each parameter. If the speed of articulatory gesture is most commonly referred to as a truly distinctive feature. We shall argue against this view, emphasizing a functional organization of the oral gestures, and propose that the sonorancy better characterizes the tap.

2. Experimental Procedures

This experiment was part of a larger study examining lingual articulations [t, d, k, k₁, q, q₁, n, p, r, 1, c, s, z, ç, ç₁] using EPG [6]. The speech items were V1CV2 disyllabic words in which the consonants [t, d, r] and the vowels /i, a, u/ were in all possible combinations. All the target words were embedded in a carrier sentence [moo_bakarida’s]. Two native speakers of Japanese, one male (MN) and one female (TM), repeated each sentence six times at self-selected normal speed, with the default accent pattern (i.e. LH pattern) on the target word. Five repetitions were analyzed. The EPG and acoustic recordings were done in the phonetics laboratory of the School of Oriental and African Studies, University of London.

The Reading EPG artificial palate has 62 electrodes arranged in 8 horizontal rows: the frontmost row (R1) has 6 electrodes and the other rows have 8 electrodes. The palatal morphology was studied from the impression of the hard palate and electrode locations were tape-measured. Three major regions were identified sagittally: front, central and back.

![Figure 1: EPG prototypical palatogram and three regions](image)

This zoning was applied to the analysis of the utterances produced by both speakers, because it was found that the corner of the alveolar ridge lay between R4 and R5 on the EPG palate of the two informants.

The configurational characteristics were studied at the point of maximum linguopalatal constriction (MAX), the EPG frame showing the maximum number of on-electrodes.

The duration of articulatory closure was measured for the temporal interval between the following two points. One is the closure onset, which was defined as the first frame showing full central contact or at least four central contacts in any one row [7]. In addition, the frame showing one or more central electrodes within the front region was regarded as a criterion for the sequences with [r] [8]. The other is the closure offset, the last frame showing complete articulatory closure.

The degree of coarticulatory sensitivity was inferred from the percentage of electrode activation presented by each target consonant at the MAX. Since the major constriction of the consonants was formed in the front region, the percentage of contact in the other regions allowed us to infer the degree of tongue dorsum elevation and the interplay between the two components of the tongue.
3. Results

3.1. Configurational Characteristics

The EPG palatograms of [t], [d], and [r] in the symmetrical contexts are given in Figure 2. The plosives [t], [d] involve the apicolabial closure along the dental-velar region with complete lateral closure. This central-lateral occlusion is always realized and stable across the vowel contexts. The constriction length is 3 rows for [t] and 2 rows for [d].

The tap [r], in contrast, involves the tongue tip (and very anterior part of the blade), forming a closure in the alveolar region. The complete closure was constantly attained by MN but not by TM: the amount of contact is rather large. Full lateral closure was not always made in the open vowel context.

3.2. Articulatory Closure Duration

Table 1 summarizes the means and standard deviations of articulatory closure duration for [t], [d], and [r]. To examine the differences between the consonants, the contextual vowels, and speakers, a three-way ANOVA was performed: 3 stops × 2 differences between the consonants, the contextual vowels, and speakers 

\[ F(2,258)=8.93, p<0.01 \]; 2 contexts × 3 stops \[ F(2,258)=5.72, p<0.01 \]; and three factors \[ F(2,258)=3.32, p<0.01 \].

A one-way ANOVA was carried out separately for the two speakers. The durational differences between the three stops were significant: for MN, \( F(2,132)=75.12, p<0.01 \); for TM, \( F(2,132)=165.78, p<0.01 \). The Newman-Keuls post-hoc analysis showed the dominant, time-related, pattern \( t > d > r \) (p<0.05) for both speakers. It is worth noting that the duration of [r] is significantly longer in MN than in TM [t=4.93, p<0.01, df=88]; and that standard deviation reveals much more variability in MN’s utterance than in TM’s. This may be related to idiosyncratic characteristics of the tip/blade gesture.

3.3. V-to-C Coarticulatory Characteristics

Figure 3 presents the mean percentage values of the Max contact for [t], [d], and [r] in the symmetrical and asymmetrical contexts. The values were calculated for the three regions separately and averaged over five repetitions. If the activity of the tongue dorsum is consonant-specifically determined, the contact variability in the posterior regions can be considered as an act of self-maintenance on the given consonantal articulation, as well as the effects of the contextual vowels.

To test the differences in the consonant type, a series of one-way ANOVA was conducted. Also, the gross effects of the changing V1s (carryover) and the changing V2s (anticipatory) were analyzed by a series of two-way ANOVA (3 V1s × 3 V2s). The results are summarized in Tables 2 and 3 respectively.

The pattern of the results in Tables 2 and 3 reveals that the amount of contact in the three regions varies systematically with the consonant type; and that co-articulatory activities vary significantly between the speakers.

Table 1: Means and standard deviations of articulatory closure duration for [t], [d], and [r] (in milliseconds) (n=270)

<table>
<thead>
<tr>
<th></th>
<th>iCi</th>
<th>aCa</th>
<th>uCu</th>
<th>iCa</th>
<th>aCi</th>
<th>iCu</th>
<th>uCi</th>
<th>aCu</th>
<th>uCa</th>
<th>pooled vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>MN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[t]</td>
<td>76.13 (4.05)</td>
<td>84.07 (3.77)</td>
<td>92.01 (5.43)</td>
<td>105.25 (18.33)</td>
<td>136.37 (13.11)</td>
<td>109.89 (11.32)</td>
<td>131.75 (10.82)</td>
<td>115.18 (2.76)</td>
<td>90.69 (7.25)</td>
<td>104.59 (21.76)</td>
</tr>
<tr>
<td>[d]</td>
<td>35.74 (2.76)</td>
<td>43.03 (9.06)</td>
<td>51.63 (6.01)</td>
<td>78.77 (13.92)</td>
<td>95.99 (7.02)</td>
<td>80.10 (4.31)</td>
<td>89.37 (7.76)</td>
<td>84.07 (5.01)</td>
<td>60.24 (5.43)</td>
<td>68.77 (21.71)</td>
</tr>
<tr>
<td>[r]</td>
<td>31.11 (6.45)</td>
<td>22.50 (2.76)</td>
<td>34.42 (8.03)</td>
<td>88.04 (17.91)</td>
<td>54.94 (10.87)</td>
<td>53.62 (9.18)</td>
<td>60.90 (8.63)</td>
<td>61.56 (13.15)</td>
<td>41.70 (5.53)</td>
<td>49.87 (21.04)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>iCi</th>
<th>aCa</th>
<th>uCu</th>
<th>iCa</th>
<th>aCi</th>
<th>iCu</th>
<th>uCi</th>
<th>aCu</th>
<th>uCa</th>
<th>pooled vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[t]</td>
<td>93.34 (18.03)</td>
<td>74.47 (1.65)</td>
<td>88.04 (8.63)</td>
<td>88.04 (7.62)</td>
<td>137.03 (19.66)</td>
<td>104.59 (10.90)</td>
<td>129.09 (18.57)</td>
<td>99.30 (4.05)</td>
<td>80.10 (4.90)</td>
<td>99.33 (23.06)</td>
</tr>
<tr>
<td>[d]</td>
<td>70.17 (2.76)</td>
<td>53.62 (1.48)</td>
<td>60.90 (3.77)</td>
<td>78.77 (4.31)</td>
<td>100.62 (14.54)</td>
<td>88.04 (4.44)</td>
<td>99.30 (6.19)</td>
<td>72.15 (4.31)</td>
<td>52.96 (3.31)</td>
<td>75.17 (18.08)</td>
</tr>
<tr>
<td>[r]</td>
<td>35.74 (3.62)</td>
<td>37.73 (1.81)</td>
<td>33.76 (3.62)</td>
<td>35.08 (2.96)</td>
<td>35.74 (5.43)</td>
<td>30.45 (5.92)</td>
<td>41.70 (2.96)</td>
<td>28.46 (1.18)</td>
<td>25.15 (6.01)</td>
<td>33.76 (6.05)</td>
</tr>
</tbody>
</table>
Figure 3: Mean percentage values of tongue-palate contact in the three regions for [t] (top), [d] (middle), and [r] (bottom)

(a) symmetrical contexts: -validate; •-validate; ▲-validate; (b) asymmetrical contexts: -validate; ▲-validate; ▲-validate; •-validate; ◻-validate

For the front region the effects of the consonant type overrule the variability of the contact amount both in the anticipatory and carryover contexts: [t] is consistently larger than [d] or [r]. There are notable coarticulatory idiosyncrasies: [t]¬-[r]-d-[d] for MN but [t]-[d]-[r] for TM. Furthermore, the effects of the flanking V1s are non-significant in MN, but are significant in TM and exhibit a reverse order, /a/¬/u/¬/t/.

For the posterior two regions, both the consonant type and the flanking vowels contribute to the variability. The patterns, /b/¬/a/-/u/ and /h/¬/a/-/t/, are dominant in the two contexts and across the consonants. However, the dorsal elevation typical for the contextual vowels differs in degree and across the speakers: [r] is the least affected of the three, yielding a general trend [t]-[d]-[r] for both speakers; but the vowel-dependent effect tends to be non-significant between [d] and [r] for MN.

These results imply that the raising gesture of the tongue dorsum is extremely antagonistic to the tongue tip/blade activity of [r]. Contrary to the claim by [4], it is more plausible to assume that the lesser amount of contact is the consequence of the synergetic interplay between the two components of the tongue. More direct evidence supporting this view is found in the temporal domain of lingualpalatal contact.

Figure 4 shows the EPG contact trajectories of the [id] and [iri] syllables. The central-region trajectory demonstrates a typical movement for each consonantal articulation. In [id] the central contact decreases immediately after the front contact attains its peak (complete closure): this can be considered as depression effects similar to the data reported by [9]. In contrast, the central trajectory in [iri] evidently reveals its substantial decrease before the front trajectory attains its peak: the vocalic and consonant gestures are altered in timing and degree. Thus, the tip/blade activity of [r] actually limits the dorsum involvement in both the spatial and temporal domains. Note in passing that the central trajectory in [iri] tends to remain constant throughout the peak of the front contact.

To examine the linear correlation between the parameters, a regression analysis was conducted. Figure 5 (a) indicates that 47% of the variation in the front region is accounted for by the closure duration: R2=0.47 [F(1,268)=243.74, p<0.01]. Further analysis between the particular pairs reveals R2=0.37 [F(1,178)=32.66, p<0.01] for [d, r]; and R2=0.25 [F(1,178)=32.66, p<0.01] for [d, r]. This weaker value comes from the durational effects of the flanking V1s are non-significant in MN, but are significant in TM and exhibit a reverse order, /a/-/u/-/t/.

Table 2: One-way ANOVA results for consonant-type effects in the three regions†

<table>
<thead>
<tr>
<th>MN</th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>/CV/</td>
<td>F(2,42)</td>
<td>F(2,42)</td>
<td>F(2,42)</td>
</tr>
<tr>
<td>/aCV/</td>
<td>F(2,42)</td>
<td>F(2,42)</td>
<td>F(2,42)</td>
</tr>
<tr>
<td>/VCu/</td>
<td>F(2,42)</td>
<td>F(2,42)</td>
<td>F(2,42)</td>
</tr>
</tbody>
</table>

Table 3: Two-way ANOVA results for anticipatory and carryover effects in the three regions of [t], [d], and [r](main effects only)†

<table>
<thead>
<tr>
<th>MN</th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/</td>
<td>F(2,36)</td>
<td>F(2,36)</td>
<td>F(2,36)</td>
</tr>
<tr>
<td>/d/</td>
<td>F(2,36)</td>
<td>F(2,36)</td>
<td>F(2,36)</td>
</tr>
<tr>
<td>/r/</td>
<td>F(2,36)</td>
<td>F(2,36)</td>
<td>F(2,36)</td>
</tr>
</tbody>
</table>

† *p<0.05; **p<0.01. The relation between the particular consonant pairs was assessed by the Newman-Keuls post-hoc analysis (p<0.05): the symbol > indicates a significant difference but the symbol = indicates non-significant difference.
properties specific to [d] and [r], since the distribution of the contact amount is in a similar range. However, the higher variability in the contact amount questions whether the speed is the most influential factor in the tap articulation.

Figure 5(b) examines intergestural coordination between the tip/blade (front) and dorsum (central). The analysis reveals the significant, but weak, R2 value: \( R^2 = 0.20 \) (\( F(1,268) = 67.22, \ p < 0.01 \)). Further analysis for each consonant shows that the R2 values are much weaker for [t] (0.03) and [r] (0.008) than for [d] (0.45). For both [t] and [r], this bare minimum correlation implies the relative independency of the two components of the tongue: also, whereas [t] apparently reflects the height of the dorsum for the vowels, [r] reflects the suppression of the vocalic gestures, a distinctive activity that might be indicative of the characteristic hollow observed by expert phoneticians. For [d], in contrast, a higher R2 value suggests that the two components of the tongue are closely interrelated.

4. Discussion

The findings of the EPG experiments are compatible with the large number of studies that have demonstrated the consonant-specific requirements for coronal stop articulations in various languages. Although the coronal stops [t, d, r] share common parameters, their controls differ systematically.

It is reasonable to assume the spatiotemporal controls of lingual gestures and configurations to be related to the aerodynamic process. The complete occlusion by the tongue tip/blade limits the size of the oral cavity and hence the level of intraoral pressure for voicing. As compared to [t], the production of [d] demonstrated a shorter constriction length, the decrease of the central trajectory, and the relatively rigid coupling between the two components of the tongue. These characteristics are supposed to be part of the expansion of the entire vocal tract to attain a large enough volume for voicing. As compared to [t], the tip/blade (front) and dorsum (central). The analysis reveals the significant, but weak, R2 value: \( R^2 = 0.20 \) (\( F(1,268) = 67.22, \ p < 0.01 \)). Further analysis for each consonant shows that the R2 values are much weaker for [t] (0.03) and [r] (0.008) than for [d] (0.45). For both [t] and [r], this bare minimum correlation implies the relative independency of the two components of the tongue: also, whereas [t] apparently reflects the height of the dorsum for the vowels, [r] reflects the suppression of the vocalic gestures, a distinctive activity that might be indicative of the characteristic hollow observed by expert phoneticians. For [d], in contrast, a higher R2 value suggests that the two components of the tongue are closely interrelated.

4. Discussion

The findings of the EPG experiments are compatible with the large number of studies that have demonstrated the consonant-specific requirements for coronal stop articulations in various languages. Although the coronal stops [t, d, r] share common parameters, their controls differ systematically.

It is reasonable to assume the spatiotemporal controls of lingual gestures and configurations to be related to the aerodynamic process. The complete occlusion by the tongue tip/blade limits the size of the oral cavity and hence the level of intraoral pressure for voicing. As compared to [t], the production of [d] demonstrated a shorter constriction length, the decrease of the central trajectory, and the relatively rigid coupling between the two components of the tongue. These characteristics are supposed to be part of the expansion of the entire vocal tract to attain a large enough volume for voicing. Thus, the lingual articulation reveals that spatiotemporal controls contribute to enhance the voicing distinction.

In contrast, it is supposed that [r] does not involve processes such as pressure build-up. It is not that the articulatory gesture is too rapid to do them, but that they are unnecessary for shaping and valving activities to generate the typical acoustic outcome. This point resolves the controversial situation in the specification of phonological features. Whereas the major class feature [-continuant] is shared by the stops [t, d, r], only the tap is co-specified by the feature [+sonorant]. The two features reflect different phases of the overall articulatory processes of [r]. Given an articulatory sequence, approach-hold-release, the release phase defines [r] as a sonorant. If we divide that phase into articulatory release

and acoustic explosion, as suggested by [10], the consonant [d] has both release and explosion (burst); in contrast, [r] has only release and no explosion after it. Furthermore, the release gesture, rather than the shorter duration of lingual closure, is relevant for creating the percept of [r] in the words such as [ʃənəri] ‘principle’. One possible source for the characteristic release is the ‘stiffness’ of the tongue dorsum that is in accordance with the observed coarticulatory behaviors. This characterization may also help us understand the fact that the tap in child’s speech is frequently deleted, or replaced by [d], and is acquired at a very later stage.

5. Conclusions

Articulatory properties of the Japanese [t, d, r] have been explored in terms of articulatory and coarticulatory patterns in the spatiotemporal domain. The findings suggest that differential properties for each stop articulation are attributed to systematic changes of the (co-)articulatory parameters. Articulatory constraints should be viewed as the functional relations to the overall state of the vocal tract, as well as the requirements for the formation and release of the given gesture.

6. References