Acoustic Characteristics of Adults’ Rhotic Monophthongs and Diphthongs

Hyunju Chung, Karen E. Pollock

Department of Communication Sciences and Disorders, Faculty of Rehabilitation Medicine, University of Alberta, Alberta, Canada

Correspondence: Hyunju Chung, PhD
Department of Communication Sciences and Disorders, Faculty of Rehabilitation Medicine, University of Alberta, 8205 114St. 2-70 Corbett Hall, Edmonton, AB, Canada, T6G 2G4.
Tel: +1-780-492-5900
Fax: +1-780-492-9333
E-mail: hyunjuchungedu@gmail.com

Received: October 27, 2013
Revised: January 14, 2014
Accepted: February 5, 2014

This study was supported by Internal Grant from the Faculty of Rehabilitation Medicine, University of Alberta.

Objectives: Rhotic sounds are known to be among the later developed sounds in young children, especially for those with speech sound disorders. Despite this fact, not many studies have examined the characteristics of rhotic vowels. Adults' productions of rhotic vowels have not been well investigated in spite of a relatively large literature on rhotic consonants. This study examined the acoustic characteristics of rhotic monophthongs and diphthongs to see if certain phonetic contexts require less demanding articulatory movements or are different in vowel duration from other contexts, and thus work as a facilitating context for the acquisition of rhotic vowels. Methods: Ten monolingual female adult speakers of Western Canadian English produced 36 target words containing two rhotic monophthongs ([ɜʰ] and [ɑ̃]), and four rhotic diphthongs (/ɜ̃ɛɹ/, /ɜ̃ɪɹ/, /ɑ̃ɛɹ/, and /ɑ̃ɪɹ/) in both open and closed syllables. Acoustic analyses were performed to extract F2 and F3 values across the vowel duration, as well as the duration for each vowel. Results: Constantly low F3-F2 values were found for rhotic monophthongs and rhotic diphthongs with pre-rhotic vowels, but steeper downward movement was found for rhotic diphthongs with back pre-rhotic vowels. Stressed and unstressed rhotic monophthongs showed similar acoustic patterns, except for [ɜʰ] which was slightly longer than [ɑ̃]. Across four rhotic diphthongs, no clear durational difference was found. Conclusion: Differences in acoustic patterns by phonetic contexts across six different rhotic vowel types suggest that certain phonetic contexts could provide more salient perceptual cues and thus facilitate relatively easier mastery of sounds over others for young children.

Keywords: Rhotic vowels, Adults, Acoustics

English rhotic sounds include the consonant [ɹ] in word-initial singletons as in words red or rock or in word-initial clusters as in words tree or frogs. These types of rhotic sounds are referred as 'consonantal /r/' or 'rhotic consonants'. Other rhotic sounds appear in the syllabic nuclei, in vocalic (vocalic /r/) or postvocalic (postvocalic /r/) position. An /r/ sound in the vocalic position is considered a 'rhotic monophthong.' Rhotic monophthongs include [ɜʰ] (stressed) as in her or bird, and [ɑ̃] (unstressed) as in tiger or zipper. Because the /r/ sounds in postvocalic position are part of the syllabic nuclei and cannot be separated from the preceding vowel sounds, they are represented as part of a diphthong (following Allen, 1979). 'Rhotic diphthongs' have [s] for the second element or offglide, as in ear ([ɨɹ]), air ([ɨɹɹ]), or ([ɜɹɹ]), or are ([ɑɹɹ]).

Acoustic characteristics, as well as articulatory configuration of rhotic sounds, have been studied extensively. The two main articulatory configurations for the production of rhotic sounds include “bunched” (tongue dorsum raised and tongue tip lowered) and “retroflex” (tongue tip raised and tongue dorsum lowered), whose supralaryngeal constrictions are made at three different locations: pharynx (narrowing), palatal vault, and at the lips (e.g., Alwan, Narayanan, & Haker, 1997; Delattre & Freeman, 1968). However, studies have shown that articulatory configurations for the production of the American English phoneme /r/ vary substantially depending on the given phonetic contexts or the speakers (e.g., De-
hyunju chung, et al.  •  Acoustic Characteristics of Adults’ Rhotic Vowels

In particular, tongue shapes are known to vary largely. Despite the different articulatory configurations required for the production of rhotic sounds, acoustic cues that mark rhoticity, however, are very stable across different phonetic contexts and speakers (e.g., Boyce & Espy-Wilson, 1997; Delattre & Freeman, 1968; Guenther et al., 1999). Generally, the F3 values of rhotic sounds are below 2,000 Hz (Hagiwara, 1995) and are often so low that they seem to be merged with F2.

These findings on the articulatory configurations and acoustic characteristics of rhotic sounds are mostly based on rhotic consonants. Some studies have included rhotic vowels, but are often limited to rhotic monophthongs; studies of rhotic diphthongs are scarce. There are a few studies that have examined one or two types of rhotic diphthongs (e.g., McGowan, Nittouer, & Manning, 2004). However, it is still not well understood whether the rhotic diphthongs, especially different types of rhotic diphthongs, show similar acoustic characteristics to rhotic consonants or rhotic monophthongs.

Understanding the characteristics of rhotic vowels as well as consonants is important for a more comprehensive understanding of rhotic sound development in young children. Studies have shown that although both rhotic vowels and consonants are known to be later developed sounds than their non-rhotic counterparts in young children (e.g., Shriberg, 1993; Smit, Hand, Freilinger, Bernthal, & Bird, 1990), rhotic vowels are known to be mastered earlier with fewer errors than rhotic consonants (Smit et al., 1990; Smit, 1993). In addition, while one of the most common error patterns for rhotic consonants is “gliding” to the [w] sound, and to a lesser extent, partial derhoticization or labialization (e.g., Smit, 1993), the pattern of errors for rhotic vowels can vary depending on the phonetic context or pre-rhotic vowel element (Pollock, 2013). This could suggest that although there is a rather consistent acoustic cue for marking rhotic sounds, the way children progress toward adult-like rhotic vowels could vary depending on the phonetic context.

This study investigated the acoustic characteristics of rhotic monophthongs (stressed and unstressed) and four different rhotic diphthongs as produced by ten female adults. F3 values, as well as F2 values, were extracted across the vowel duration to examine changes in formant values over time. The term “vowel duration” in this paper indicates the duration of any pre-rhotic element as well as the rhotic element of each vowel. In addition, slopes of both F3 and F3-F2 were also examined to understand if certain phonetic contexts require more changes in vocal tract configuration than others. Lastly, vowel duration was also investigated across the different target vowels.

Specific research questions included:
Are there differences in vowel duration and acoustic characteristics of female adults’ production of rhotic vowels:
1) By complexity (monophthongs vs. diphthongs);
2) By stress pattern (stressed vs. unstressed);
3) By pre-rhotic vowel element of rhotic diphthongs (front vs. back vowels);
4) By syllable structure (open vs. closed syllables) for both rhotic monophthongs and diphthongs.

METHODS

Participants
Participants included 10 monolingual female adult speakers of Western Canadian English. The age of participants ranged from 20 to 36 years (mean age 26, SD 5.32). All of the participants were born in Western Canada (British Columbia or Alberta) and had spent all or most of their lives in the same region. All reported no history of speech, language, or hearing problems.

Stimuli
Targets included 36 words containing rhotic monophthong /ɜ/ and its unstressed variant [ɻ], and four rhotic diphthongs (/ɪɿ, ɛɿ, ɔɿ, ɑɿ/) in both open and closed syllables. The full list of target words is provided in Table 1.

Procedure
The speech samples were collected in a quiet room at the Uni-

<table>
<thead>
<tr>
<th>Table 1. List of target words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Monophthongs</strong></td>
</tr>
<tr>
<td>ɜ, ŋ</td>
</tr>
<tr>
<td>ŋ</td>
</tr>
<tr>
<td><strong>Diphthongs</strong></td>
</tr>
<tr>
<td>ɪɿ, ɜɿ</td>
</tr>
<tr>
<td>ɛɿ</td>
</tr>
<tr>
<td>ɔɿ</td>
</tr>
<tr>
<td>ɑɿ</td>
</tr>
</tbody>
</table>
versity of Alberta. Target words were elicited using a picture naming task. Participants were asked to spontaneously produce each target word after viewing color pictures presented on a laptop screen. Each target word was produced once for each speaker. Productions were recorded using a Shure WH20 head mounted microphone and a Marantz Professional PMD661 high quality audio recorder at a sampling rate of 44.1 kHz and with 16-bit quantization.

Acoustic analyses

The onset and offset of each target vowel were determined using the acoustic analysis software Praat (version 5.3.16; Boersma & Weenink, 2012). The vowel onset was marked where a systematic waveform and a clear glottal pulse started to appear, and the vowel offset was marked where a systematic waveform and the second formant frequency (F2) started to fade out (Figure 1). Once the vowel boundary was defined, the formant measurements of each vowel were obtained using a custom speech analysis program (Morrison & Nearey, 2011) created in MATLAB version 7.8.0.347 (R2009a). This program gives 8 different formant tracking options for each vowel to be analyzed. Among the options, the first three formant frequencies (F1, F2, and F3) that best represent the actual production were extracted for each 2 ms step over the entire duration of the vowel. Minor mistrackings were then smoothed manually. Values from the first 10 ms were excluded to minimize the coarticulatory effect of word-initial consonants. That is, the onset values were extracted at the 12 ms time point and the offset values at the 12 ms before the last time point. For each token, F3 slope was calculated by dividing the difference between F3 at the 12 ms time point and at the 12 ms before the last time point by the duration of each vowel (subtracting 12 ms from the time point that was 12 ms before the last time point). F3-F2 slope was also calculated using the same formula to examine the rate of change in F3-F2 values. Among the total of 360 tokens (36 words × 10 adults) produced, 7 tokens were excluded due to poor sound quality or background noise.

RESULTS

The set of acoustic values of the six target rhotic vowels are provided in Table 2. This includes the average minimum F3, F3, and

![Figure 1. Praat measurements of the onset (upper) and offset (bottom) of the rhotic vowel /ɪɚ/ as in beard.](image)

<table>
<thead>
<tr>
<th></th>
<th>Monophthongs</th>
<th>Diphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ɪɚ</td>
<td>ɪɚ</td>
</tr>
<tr>
<td>minF3</td>
<td>1,787</td>
<td>1,925</td>
</tr>
<tr>
<td>F3 onset</td>
<td>2,127</td>
<td>2,375</td>
</tr>
<tr>
<td>F3 offset</td>
<td>2,144</td>
<td>2,419</td>
</tr>
<tr>
<td>F3-F2 onset</td>
<td>498</td>
<td>595</td>
</tr>
<tr>
<td>F3-F2 offset</td>
<td>344</td>
<td>505</td>
</tr>
<tr>
<td>F3 slope</td>
<td>.19 (1.59)</td>
<td>.22 (2.96)</td>
</tr>
<tr>
<td>F3-F2 slope</td>
<td>-.63 (1.11)</td>
<td>-.55 (2.61)</td>
</tr>
</tbody>
</table>

Table 2. Average minimum F3, F3 and F3-F2 (Hz) at the vowel onset and offset, and means and standard deviations (in parenthesis) of F3 and F3-F2 slopes for the six target rhotic vowels.
F3-F2 (Hz) at the vowel onset and offset, and the mean and standard deviation of F3 and F3-F2 slopes for the six target rhotic vowels. F2 and F3 values over the duration of the vowel are depicted in Figure 2. As can be observed in Figure 2, the F3 of all six rhotic vowels either reached or were very close to the average F3 value of /ɜ/ (1,870 Hz) for female adults reported in Assmann & Katz (2000), confirming low F3 values of rhotic vowels. Both stressed and unstressed rhotic monophthongs had stable F3 values across the vowel duration, while those of rhotic diphthongs had relatively clear F3 downward movements in both open and closed syllables. Rhotic diphthongs with front pre-rhotic vowels (ɪɚ/ and ɛɚ/) had steeper downward movement (mean F3 slope ɪɚ/ -3.69; ɛɚ/ -2.29) than those with back pre-rhotic vowels (ɔɚ/ -1.81; ɑɚ/ -1.29). As confirmed by one-way ANOVA, F3 slopes differed significantly by rhotic vowel type (F5,347 = 32.59, p < .05), and a Tukey HSD post-hoc test showed that F3 slopes of each rhotic vowel differed significantly from each other, except for the [ɪ]-[ɜ]/, ɪɚ/-ɛɚ/, and ɪɚ/-ɜɚ/ pairs. Between open and closed syllables, a Welch two-sample t-test showed that F3 slopes differed significantly by syllable structure in all rhotic vowels except for ɪɚ/. The output is summarized in Table 3.

Figure 3 shows the value of F3-F2 across the duration of the vowel. The F3-F2 values of the two rhotic diphthongs with front pre-rhotic sounds, ɪɚ/ and ɛɚ/, showed constant low F3-F2 values across the full vowel duration, similar to those of the rhotic monophthongs. This contrasting pattern between F3 and F3-F2 patterns of ɪɚ/ and ɛɚ/ is due to the pattern of F2, which shows a similar downward shape as F3, thus resulting in constant low values for F3-F2. In contrast, F3-F2 values of rhotic diphthongs with back pre-rhotic vowels, ɔɚ/ and ɑɚ/, had relatively steep F3 downward movement (slope of ɔɚ/ 3.98; ɑɚ/ -2.28), which did not reach the average F3-F2 values of /ɜ/ (362 Hz) from Assmann & Katz (2000) until the end of the vowel duration. The diphthong ɔɚ/ showed the steepest movement, where the difference between the average onset (1,265 Hz) and offset (378 Hz) F3-F2 value was as big as 887 Hz. Although the two rhotic diphthongs with front pre-rhotic sounds had similar patterns of F3-F2 movements and F3-F2 slopes as the rhotic monophthongs, the F3-F2 slopes of ɪɚ/ and ɛɚ/ were slightly steeper than those of the monophthongs. In addition, ɪɚ/ and ɛɚ/ were longer in duration than rhotic monophthongs. That is, as ex-

Table 3. Output of the Welch two-sample t-test for F3 slope by syllable structure (open vs. closed syllable) for each rhotic vowel

<table>
<thead>
<tr>
<th></th>
<th>ɪɚ/</th>
<th>ɜɚ/</th>
<th>ɪɚ/</th>
<th>ɛɚ/</th>
<th>ɔɚ/</th>
<th>ɑɚ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-value</td>
<td>-4.171</td>
<td>-4.4377</td>
<td>-3.8117</td>
<td>-2.2863</td>
<td>0.1387</td>
<td>-2.9517</td>
</tr>
<tr>
<td>df</td>
<td>58</td>
<td>57</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>p-value</td>
<td>.0001</td>
<td>4.216e-05</td>
<td>.0003</td>
<td>.0260</td>
<td>.8902</td>
<td>.0046</td>
</tr>
</tbody>
</table>
expected, rhotic diphthongs were in general longer than monophthongs. Between the two types of rhotic monophthongs, [ɝ] (unstressed) was significantly shorter than its stressed counterpart [ɚ]. However, the duration of vowels in closed syllables was influenced by the voicing of the final consonants, which usually yield longer vowels than those followed by voiceless word final consonants, as shown in Table 4. The mean and standard deviation of vowel duration for each rhotic vowel in both open and closed syllables are summarized in Table 4.

In summary, it was found that the characteristics of rhotic vowels differ by phonetic context: by complexity (monophthongs vs. diphthongs), by stress pattern ([ɝ] vs. [ɚ]), by pre-rhotic vowel element of rhotic diphthongs (front vs. back vowels), and by syllable structure (open vs. closed). Rhotic monophthongs showed more stable F3 slopes than rhotic diphthongs and were significantly shorter in duration than rhotic diphthongs. For the two rhotic monophthongs that differed in stress pattern, stressed [ɝ] was significantly longer than unstressed [ɚ], although no significant difference in their F3 slope pattern was found. Rhotic diphthongs with front pre-rhotic vowel elements were significantly longer in duration and had significantly steeper F3 slope than those with back pre-rhotic vowel elements.

### Table 4. Means and standard deviations (in parenthesis) of vowel duration (ms) of six rhotic vowels in both open and closed syllables

<table>
<thead>
<tr>
<th></th>
<th>Monophthongs</th>
<th>Diphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ɝ</td>
<td>ɚ</td>
</tr>
<tr>
<td>Open syllable</td>
<td>259.89 (62.57)</td>
<td>172.60 (46.62)</td>
</tr>
<tr>
<td>Closed syllable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voiced coda</td>
<td>249.26 (66.96)</td>
<td>192.55 (28.27)</td>
</tr>
<tr>
<td>Voiceless coda</td>
<td>139.06 (25.38)</td>
<td>NA^a</td>
</tr>
</tbody>
</table>

^a All of the target words for [Ʉ] were followed by voiceless consonants.

^b All of the target words for [Ʉ], [Ʉɪ̆], and [Ʉɛ̆] were followed by voiced consonants.

### CONCLUSION

This study examined the acoustic characteristics of rhotic vowels produced by ten female adults. The acoustic characteristics of six different target rhotic vowels were compared to each other by complexity (monophthongs vs. diphthongs), by stress pattern ([ɝ] vs. [ɚ]), by pre-rhotic vowels (front vs. back), and by syllable structure (open vs. closed).

The analysis of F3-F2 movement patterns revealed constantly low F3-F2 values for rhotic monophthongs and rhotic diphthongs with front pre-rhotic vowels. Rhotic diphthongs with back pre-rhotic vowels, on the other hand, showed a steeper downward movement. No clear difference in F3-F2 slope was found between rhotic monophthongs and rhotic diphthongs with front pre-rhotic vowels, although [Ʉɪ̆r] and [Ʉɛ̆r] had slightly steeper slopes than those of rhotic monophthongs. Given that the smaller distance between F2 and F3 is the main cue for rhoticity, this pattern could suggest that rhotic monophthongs and rhotic diphthongs with front pre-rhotic vowels could have more perceptually salient cues for rhoticity than those with back pre-rhotic vowels, especially [Ʉɛ̆r], which had the steepest downward movement. That is, rhotic diphthongs with back pre-rhotic vowels could be more challenging for young chil-
children to master than rhotic monophthongs or rhotic diphthongs with front pre-rhotic vowels, since the acoustic cue for rhoticity could not be achieved until the end of the vowel.

No clear difference in F3-F2 or F3 slope was found between stressed and unstressed rhotic monophthongs. The only apparent difference between the two vowels was the duration, with [ɾ] being longer than [ɾ]. This could mean that for the development of rhotic monophthongs, once young children get the target F3 (low F3 value), having the right durational difference between [ɾ] and [ɾ] is important for having adult-like production of [ɾ] and [ɾ].

Although further analysis with a larger data set could be more informative, we can offer some initial observations about the effect of phonetic context on the production of rhotic vowels. For example, the results of this study suggest that rhotic vowels in certain phonetic contexts could be easier to produce than those in other contexts that require steeper formant movement (thus, rapid changes in vocal tract configuration) or are shorter in duration. This information could be particularly informative in understanding rhotic vowel development in young children, especially children with speech sound disorders, who usually experience difficulty in producing rhotic sounds. Given that rhotic sounds are among the most challenging sounds for children to master, information on facilitating phonetic contexts could be useful in selecting specific rhotic vowels to be targeted first in intervention. In addition, second language learners of English, especially those whose native language does not have English-like /ɾ/ phonemes (e.g., Korean), could use this information to target their rhotic vowel productions. Lastly, these systematic analyses of different rhotic vowels could serve as reference data for understanding children’s production of rhotic vowels. Further research will be carried out to evaluate the characteristics and development of rhotic vowels in young children.

REFERENCES


국문초록

성인 치경접근 단모음과 이중모음의 음향학적 특징
정현주·Karen E. Pollock
캐나다 앨버타대학교 언어병리학과


핵심어: /r/ 모음 (치경 접근모음), 성인발화, 음향학
본 연구는 캐나다 앨버타 대학교 재활의학 단대 내의 연구비를 지원받아 연구되었음.