LIP POSTURES OF HIGH VOWELS IN TAIWAN MANDARIN

Chenhao Chiu[1, 2], Po-Hsuan Huang[1]

Graduate Institute of Linguistics, National Taiwan University[1]
Neurobiology and Cognitive Science Center, National Taiwan University[2]
chenhaochiu@ntu.edu.tw, r09142003@ntu.edu.tw

ABSTRACT

Taiwan Mandarin contrasts three high vowels with frontness ([i] and [y] being front) and roundedness ([y] and [u] being rounded). In particular, the [round] feature appears to provide rather simplistic interpretations of the lip postures for these high vowels as this feature is associated more with a narrowed aperture than with lip posture. Whether postural differences are articulatorily achieved to distinguish these high vowels from one another remains unanswered. This study investigates the lip postures among high vowels in Taiwan Mandarin, including aperture distance and area, axial ratio, and lip protrusion. Our results show that [u] and [y], though both traditionally labeled as rounded, contrast with each other in lateral distance between the mouth corners, yielding a more circular round posture for [u] and a more laterally compressed posture for [y]. Collectively, our results suggest that high vowels in Taiwan Mandarin are better distinguished along aperture area and lip posture.

Keywords: lip postures, lip aperture, roundedness, high vowels, Taiwan Mandarin.

1. INTRODUCTION

The vowel system in Mandarin constitutes a three-way contrast: [i], [y], and [u] ([3, 5]). These three vowels are distinguished from each other on two dimensions: frontness and roundedness. These two dimensions are associated with two distinctive features, [back] and [round], in phonology term. The high front vowel [i] is assigned with [−back] and [−round], [y] is associated with [−back] but [+round], and [u] is [+back] and [+round]. With these feature assignments, it is suggested that sounds with the same feature value share the same articulatory gestures. For example, [y] and [u] are both assigned with a positive value for the [round] feature, implying that the roundedness for these two sounds is equivocal. However, whether or not these two sounds are identical in terms of their lip postures is not determined or specified by the assignment of their feature values. The +/− values only provide dichotomic categories between sounds; either all or nothing, and nothing in between. Chances are the degree differences with regards to the roundedness is not fully revealed.

In addition, the [round] feature is resulted from two articulatory gestures: lip rounding and lip protrusion. While these two articulatory gestures can be independently controlled, the co-occurrence between them are far from uncommon, which consequently leads to the interpretation that the [round] feature is the combination of lip rounding and lip protrusion. On the other hand, the high front vowel [i] is assigned with a negative value for the [round] feature. The connotation of [−round] is associated with lips not being rounded. How the lips are actually postured is not specified; the lips can be relaxed (as in [ə]), wide open (as in [a]), or bilaterally pulled (as in [i]), etc. In a broader sense, the [−round] feature fails to capture the fact that the lips are pulled bilaterally for [i] such that the lip posture for [i] is suggested to be comparable to those for [a] and [ə] since they are all assigned with the same [−round] feature. As such, the [round] feature appears to provide rather simplistic interpretations of the lip postures for these high vowels. Whether or not subtle differences can be articulatorily achieved to distinguish these high vowels from one another remains unanswered. Phonological feature assignments may not be able to fully characterize how the lips of Mandarin high vowels are articulatorily postured.

With regards to lip rounding, protrusion and vertical compression are the two common parameters to quantify how labial sounds are produced or acquired [3, 8]. These two parameters, however, only characterize two dimensions, namely vertical and depth, out of a 3D space. The horizontal information, that is, how lips are laterally spread, and the aperture area enclosed by the vertical and lateral axes not only provide ample visual cues but also contribute to the generated acoustics. More crucially, labial sounds do not always have the same degrees or types of rounding. Catford [2] identified two types of rounding: endolabial (e.g., [u] or [o]) and exolabial (e.g., [y] or [ø]); the former refers to the rounding of the inner circle of the lips and the latter is associated with the outer rounding of the lips. Endolabial and exolabial sounds contrast each other in two aspects. First, endolabial sounds pull the corners of the mouth close together whereas exolabial sounds do not.

Second, endolabial sounds are usually accompanied by obvious lip protrusion whereas the degree of lip protrusion is more restrained for exolabial sounds [1].

While [i] – [y] differentiation in Mandarin is argued to lie in vertical distance between the lips [7], whether the lateral distance (the distance between the corners of the mouth) may also play a role in differentiating high vowels appears unanswered.

The current study examines if there is any postural difference among the high vowels in Mandarin and whether or not these subtle differences are consistently realized in production.

2. METHODS

2.1. Participants

Eighteen native speakers of Taiwan Mandarin (9 F, mean age = 23.44) participated in the experiment. None of them reported any auditory or visual disabilities. The study was conducted in accordance with ethical guidelines approved by XXX University. All participants were compensated monetarily for their time.

2.2. Apparatus

An iPad Pro (2021, 12.9 inch) was set up in front of participants (at a distance of approximately 0.5m) to capture their lip movements and, at 30 fps in 4k resolution. Four melting beads were attached to the following locations: the nose, the center of the upper vermillion boundary, the center of the lower vermillion boundary, and the right corner of the mouth (Figure 1). These beads were positioned in order to calculate the lip protrusion and lip aperture. A mirror (7 × 5.5 inch) was placed behind the participants at 45° off the midline so that the melting beads can be captured by the camera and measurement of lip aperture and protrusion and be quantified.

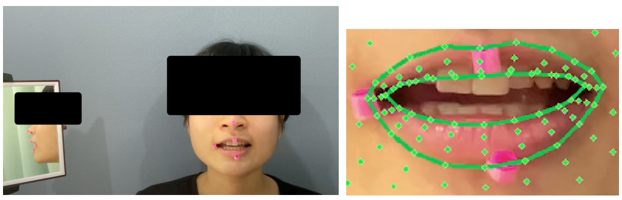

Figure 1: Marker placements for the quantification of lip protrusion and aperture (left), and MediaPipe landmarks on the lips (right).

2.3. Stimuli and procedures

Critical stimuli consisted of the three high vowels in Taiwan Mandarin (i.e., [i], [y], and [u]), produced both in isolation (i.e., monosyllabic) and embedded in disyllabic words. Monosyllabic words were matched with a high level tone (Tone 1): [i1] (‘*one*’), [u1] (‘*house*’), and [y1] (‘*mud*’). Disyllabic words all carried a low-dipping tone (Tone 3), following the same adjective with a falling tone (Tone 4): [ta4 i3], (‘*big ant*’), [ta4 u3] (‘*fifth-grader in college*’), and [ta4 y3], (‘*heavy rain*’). The experiment was designed as a self-paced reading task. Each word was presented via Microsoft PowerPoint, with ten repetitions each.

2.4. Measurements and data analyses

Face detection and tracking were performed by MediaPipe Face Mesh (powered by Google: https://google.github.io/mediapipe/solutions/face_mesh.html). MediaPipe Face Mesh provides 3D information of the 468 landmarks on the face, including the corners of the mouth and the vermillion borders of the upper and lower lips. With these landmarks, four measurements were conducted:

1. *Lateral distance*: the distance between the two corners of the inner lip borders.
2. *Vertical distance*: the maximum distance between the upper and lower lips.
3. *Axial ratio*: the ratio between lateral (numerator) and vertical (denominator) distances, measured from the inner lip borders.
4. *Aperture area*: the area enclosed by the (inner) lip borders.

The degree of lip protrusion was quantified by subtracting the position of the nose from the lips. With the relative positions of the melting beads, another two measurements were performed:

5. *Upper lip protrusion*: the relative positions between the upper lip and nose.
6. *Lower lip protrusion*: the relative positions between the lower lip and nose.

To characterize the overall lip postures, *kinematic stiffness* of the lips was also calculated, following [4, 6]. Multiple MediaPipe landmarks for the upper and lower lips were selected for the quantification of stiffness. To visualize the differences of stiffness between the lip center and corners, data were fitted through the Generalized Additive Mixed Models (GAMMs, [9]).

All measurements were then submitted to linear mixed models (LMM); with each measurement as dependent variable, and VOWEL ([i], [y], [u]) and WORD (isolation vs. disyllabic word) as fixed effects. Random slopes for VOWEL and WORD and random intercept for participant were also included. For all the model constructions, the monosyllabic [y] was set as the reference. For the model of stiffness, we also included LANDMARK (from center to corner) as fixed effects, and monosyllables and disyllables were analyzed through two separate models. All measurements were first z-transformed before analyses. Significance levels were determined at .05.

3. RESULTS

First, the results showed that the lateral distance for monosyllabic [y] was significantly shorter than that for monosyllabic [i] (β = 1.23, p <.01) but significantly longer than monosyllabic [u] ($\beta = -0.34$, p =.01). The shortest distance was found in the production of [u], followed by [y], and then [i] (Figure 2). Less lateral pull (i.e., larger lateral distance) was observed in disyllabic words than in monosyllabic words and this pattern is across-the-broad for all three high vowels, though significant differences between monosyllabic and disyllabic words were found in [y] ($\beta = 0.51$, p <.01) and [i], as indicated by the absence of a two-way interaction (β = 0.03, p =0.63).

Jackson and McGowan [7] showed that Mandarin [i] and [y] distinction lies in vertical distance. Our results (Figure 3) echo their findings in that compared with monosyllabic [y], the vertical distance for monosyllabic [i] was significantly larger (β = 1.18, p <.01). No difference was found between monosyllabic [y] and [u] (p =.72). Shorter vertical distances were observed for disyllabic [y] and [u] than their monosyllabic counterparts ([y]: β = 0.33, p =.04; [u]: β = 0.15, p =.09), but the opposite pattern was reported for [i] (β = 0.32, p <.01).

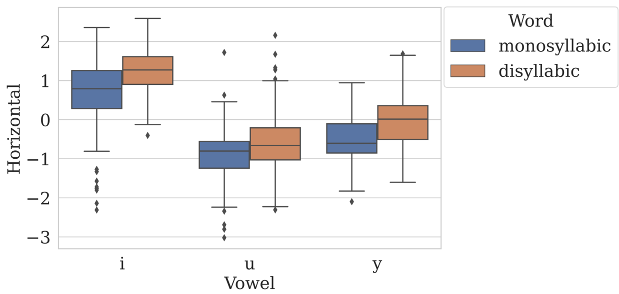

Figure 2: Lateral distance (z-scored) by vowels and conditions.

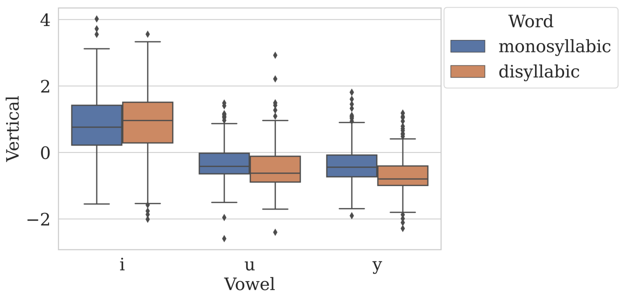

Figure 3: Vertical distance (z-scored) by vowels and conditions.

The results of axial ratios showed that the ratio for monosyllabic [y] were significantly larger than that for [i] (β = −0.81, p =. 01) while the ratios for monosyllabic [y] and [u] did not differ from each other (p = .27). Significant difference between monosyllabic and disyllabic words only lied in [y] (β = 0.56, p <. 01), but not in the other two vowels, as suggested by the VOWEL × WORD interactions ([i]: $\beta = -0.54$, p <.01; [u]: $\beta = -0.26$, p =.01).

The results of lip aperture area are visualized in Figure 3. The area of monosyllabic words for [y] was significantly smaller than that for [i] ($\beta = 1.42$, p <.01) but did not differ from [u] (p =.58). A difference between monosyllabic and disyllabic words was only found in [i], suggested by the significant VOWEL × WORD interaction (β = 0.23, p <.01), but not in the other two vowels ([y]: $\beta = -0.13$, p =.25; [u]: $\beta = 0.09$, p =.24).

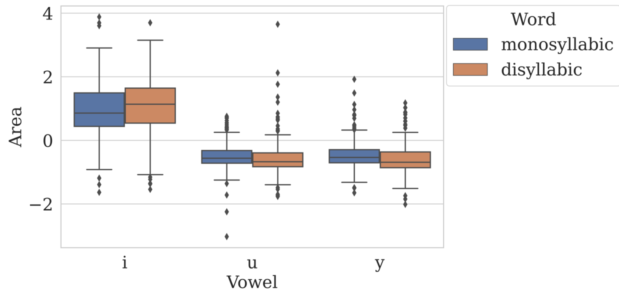

Figure 3: Lip aperture (z-scored) by vowels and conditions.

Considering lip protrusion, the upper lip was more protruded for monosyllabic [y] than for monosyllabic [i] ($\beta = -1.11$, p <.01) whereas the protrusion degree did not differ between [y] and [u] (p =.65). While mono and disyllabic [y] did not differ in the degree of lip protrusion (p =.43), interactions between VOWEL and WORD were reported for the other two vowels ([i]: $\beta = -0.21$, p =.02; [u]: $\beta = 0.33$, p <.01). Similar patterns were also observed in the lower lip protrusion: more protrusion in monosyllabic [y] than in [i] counterpart ($\beta = -0.91$, p <.01) and no difference between [y] and [u] (p =.09). No difference between mono and disyllabic [y] and between mono and disyllabic [u] (p =.66, p = .99, respectively), but a VOWEL × WORD interaction for [i] was observed (β = −0.54, p =.02). Results of upper and lower lip protrusion are presented in Figures 4 and 5, respectively.

With regards to the kinematic stiffness, the GAMMs results revealed that for all three vowels, the stiffness of the lip center was higher than those of the corners (Figure 6). This tendency was observed in both mono and di-syllables. The LMM results for monosyllables failed to find significant difference between [y] and [i] (p = .51) and between [y] and [u] (p = .13); the results for disyllables only reported a marginal difference between [y] and [i] (p =.09) but no difference between [y] and [u] (p =.67).

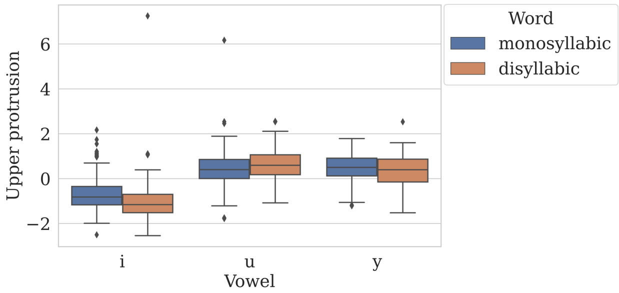

Figure 4: Upper lip protrusion (z-scored) across vowels and conditions.

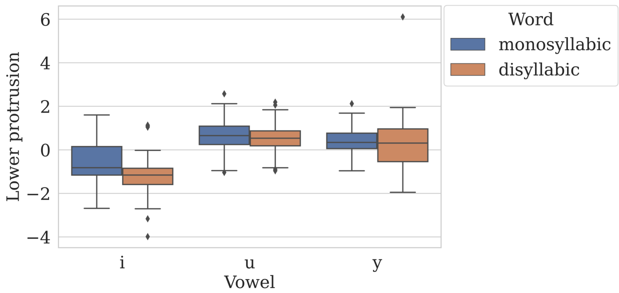

Figure 5: Lower lip protrusion (z-scored) across vowels and conditions.

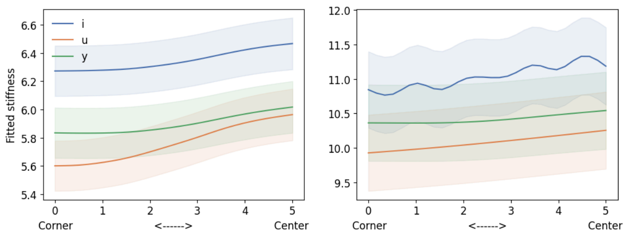

Figure 6: GAMMs predicted fits of kinematic stiffness across different vowels in mono-syllables (left) and di-syllables (right).

4. DISCUSSION

The three vowels in Taiwan Mandarin are traditionally identified through their frontness and roundedness, assuming that segments associated with the same dichotomic value would bear the same articulatory posture. In this vein, the assignment of [+round] for both [y] and [u] leads to an interpretation that the roundedness for these two sounds would be of the same degree. The current study examines the lip postures of the three high vowels in Taiwan Mandarin and compared in what articulatory dimensions they contrast with one another.

Extended from the findings in Jackson and McGowan [7], our results show that the high front vowel [i] is postured significantly different from the other two vowels in all six measurements employed; longer vertical and lateral distances, smaller axial ratios, larger lip aperture, and less lip protrusion. Longer distance both vertically and laterally along with larger lip aperture collectively point to a spread and open lip posture for the high front vowel [i]. These dimensions provide the speakers with visually distinctive cues to tell it from the other two high vowels.

The two “rounded” vowels, [y] and [u], both required shorter lip distances, smaller lip aperture, and more pronounced lip protrusion. According to Catford [2], endolabial [u] and exolabial [y] bear different degrees of lip protrusion: more pronounced for [u] than for [y]. In our results, [u] and [y] exhibited comparable lip protrusion for both upper and lower lips, suggesting that in contrast to the postural difference at the corners of the mouth, the [round] feature may be more associated with lip protrusion since both [u] and [y] are assigned with this feature.

Meanwhile, it is noted that Taiwan Mandarin [y] and [u] contrast with each other only in lateral distances, but not in other measurements. Compared with [u], larger lateral distance along with a larger axial ratio for [y] yield a lip aperture of ellipsis, echoing Catford’s [2] descriptions about the two labial sounds. The [round] feature may not fully capture this postural difference between these two vowels. Despite the fact that [y] and [u] in Taiwan Mandarin only differ in lateral distance, the yielded lip postures can be characterized by other measurements, such as the kinematic stiffness. Our stiffness results demonstrated that while both [y] and [u] had stronger stiffness for the lip center than corners, in general the lip postures for [y] were of slightly stronger stiffness than [u], particularly for the corners of the lips (Figure 6), suggesting that [y] and [u] may be differentiated from each other at the corners of the lips. Additionally, the reported stiffness changes over the lips across the three vowels are not only visually identified but also are associated with physiological accounts. Mayer et al. [10] used 3D face model to simulate lip spreading for [i] and lip compression for [u]. Their results suggest that these lip postures require different groups of muscles while quantal properties are exhibited. Collectively, these observed postural differences for the three high vowels in Taiwan Mandarin ought to be available to the speakers and therefore may serve its function for perceptual identification.

The current study compared perioral postures for the three high vowels in Taiwan Mandarin. While lip aperture characterizes the high front vowel [i], the difference between [u] and [y] did not reside in the degree of roundedness or protrusion, but rather in the postural difference at the corners of the mouth. The characterizations and measurements of these areas would call for future research.

5. REFERENCES

[1] Linker, W. 1982. Articulatory and acoustic correlates of labial activity in vowels: A cross-linguistic study. *UCLA Working Papers in Phonetics* 56.

[2] Catford, J. C. 1988. *A Practical Introduction to Phonetics*. Oxford: Clarendon Press.

[3] Ladefoged, P., Maddieson, I. 1996. *The Sounds of the World's Languages*. Oxford, England: Blackwell.

[4] Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. 2002. Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *J. Acous. Soc. Am.* 112(4), 1627-1641.

[5] Lin, Y. H. 2007. *The Sounds of Chinese.* Cambridge University Press.

[6] Van Lieshout, P. H., Bose, A., Square, P. A., & Steele, C. M. 2007. Speech motor control in fluent and dysfluent speech production of an individual with apraxia of speech and Broca's aphasia. *Clinical Linguistics & Phonetics* 21(3), 159-188.

[7] Jackson, M. T.-T., McGowan, R. S. 2012. A study of high front vowels with articulatory data and acoustic simulations. *J. Acoust. Soc. Am.* 131, 3017–3035.

[8] Saito, H. 2016. Lip movements for an unfamiliar vowel: Mandarin front rounded vowel produced by Japanese speakers. *JSLHR* 59(6), S1558-S1565.

[9] Wieling, M. 2018. Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *J. of Phon.* 70, 86-116.

[10] Mayer, C., Chiu, C., Gick, B. 2021. Biomechanical simulation of lip compression and spreading. *Canadian Acoustics 49*(3), 38 – 39.