



# A real-time MRI study on asymmetry in velum dynamics during VCV production with nasal sounds

Chetan Sharma<sup>1</sup>, Vaishnavi Chandwanshi<sup>2</sup>, Shreya Shrikant Karkun<sup>1</sup>, Aditya Anand Gupta<sup>1</sup>,  
Prasanta Kumar Ghosh<sup>1</sup>

<sup>1</sup>Electrical Engineering, Indian Institute of Science (IISc), Bangalore-560012, India

<sup>2</sup>Rewa Engineering College, Rewa, Madhya Pradesh-486002, India

chetansharma@iisc.ac.in, vcchandwanshi@gmail.com, shreyas1@iisc.ac.in, adityaag@iisc.ac.in,  
prasantg@iisc.ac.in

## Abstract

Velum movement controls airflow through nasal passage enabling production of nasal sound. The asymmetry in velum dynamics during velum lowering vs raising is not well understood, although several studies on asymmetry of other articulators exist. In this study of asymmetry in velum dynamics, we use real-time MRI videos of 68 speakers speaking symmetric vowel-consonant-vowel (VCV) sequence with V being vowels (/a/, /i/, /u/) and C being nasals (/m/, /n/). The asymmetry is analyzed in terms of the extent and speed of velum movement. The study reveals that the extent of velum displacement is significantly higher (by a factor of  $\sim 2$ ) during V-C transition (velum lowering) compared to that during C-V transition (velum raising) for all six vowel and nasal combinations chosen. It is also found that the speed with which velum lowers is higher ( $\sim 1.5$  times) than that of velum raising during production of all VCVs except /ana/ and /unu/, for which no significant difference is observed.

**Index Terms:** nasal consonants, asymmetry, vowel-consonant-vowel (VCV)

## 1. Introduction

The velum dynamics play a crucial role in speech production, linguistic variation, and clinical diagnosis. The velum actively regulates the balance between oral and nasal airflow, shaping speech sounds and contributing to phonetic and phonological contrasts across languages. Velum dynamics are known to contribute to coarticulation and sound change, such as nasalization's effect on consonant duration [1, 2]. Understanding velopharyngeal (VP) variability aids in treatment [3] and helps distinguish dysfunction from speech rate effects [4]. Velar movement estimation is also valuable for speech recognition, enhancement, and speaker identification [5, 6]. In speech synthesis, modeling velar movement enhances articulatory synthesizers and improves voice quality prediction, especially for contrastive nasalisation [7, 8]. Beyond linguistic and clinical perspectives, velum dynamics influence sociophonetics, linking lexical and social factors to phonetic variation [9, 10]. Thus, studying velum dynamics provides insights across phonetics, speech pathology, speech technology, and language evolution.

A large number of studies (a few among which are summarised in Table 1) have been carried out in the past on different aspects of velum dynamics using data of various modalities such as real-time magnetic resonance imaging (rtMRI), X-ray, and electromagnetic articulography (EMA). Works that study velum dynamics (marked blue in Table 1) address various characteristics of velum, e.g., area, shape, thickness, height and

The authors thank the Department of Science and Technology (DST), Govt of India for their support in this work.

Table 1: Summary of different studies on velum dynamics from the literature. Rows marked red indicate studies related to static properties of the velum, while those marked blue indicate studies on velum dynamics (F=Female, M=Male).

Reference(Year)	Study	Modality	Subjects	Stimuli
Kotlarek et al[11] (2019)	Velum Shape	MRI	85 children, 85 adults	-
Joshua M. Inouye et al[3] (2015)	Velopharyngeal (VP) closure force	MRI	10M	3D Random anatomies
Ibrahim Korhan et al[12] (2015)	Uvula width and length	Endoscopy	44M + 36F	-
Jianwu Dang et al[13] (1996)	Nasal Coupling	MRI	3M	Japanese vowels
Sishi Liao et al[14] (2024)	Velum movement	rtMRI	3M + 4F	CGVN /(G)an/ and /(G)al
Annabelle Purnomo et al[15] (2022)	Velum Velocity	X-Ray	5M + 4F	17 sentences
Miran Oh et al[16] (2021)	Velum position and duration	rtMRI	5 Korean speakers	(/n/), (/np/), (/nt/), (/nn/)
Jianwu Dang et al[17] (2016)	Velum thickness	rtMRI	3 Japanese speakers	/bibi/ and /bobob/
Barlaz et al[18] (2015)	Velum constriction	rtMRI	1F	Consonants /n/ and /m/
João Freitas et al[19] (2014)	Velopharyngeal area	rtMRI and EMG	3F	EP nasal vowels
Rossato et al[20] (2003)	Velum height	EMA	1M	VCV
Kuehn et al[21] (1998)	Velopharyngeal (VP) closure force	EMG	7M + 7F	/VCV/, /VNV/, /VCNV/, /VNCV/
Dang et al[22] (1994)	Velum vibration effects	X-Ray	2M	/m/ and /n/

velocity, while producing nasal sounds through different types of speech stimuli. Most of these velum dynamics studies use very few subjects due to potential challenges of recording velum movement and reliability of annotation.

Although numerous studies have investigated velum dynamics in the past, the asymmetry in velum dynamics (over time) in the VCV (C being nasal sound) production is not well explored. However, the analysis of asymmetry in movement of other articulators has been an area of research for many years. Various studies have examined the asymmetry in different contexts, for instance, identifying the asymmetry across tongue-palate contact, the asymmetry of the tongue on the development of jaws and the position of teeth, and many more. Asymmetry in tongue-palate contact was analysed using electropalatography to measure lateral asymmetry in British English linguopalatal consonants [23]. Similarly, an analysis of 1,502 palatogram data confirmed that 83% of speech sounds show asymmetry, occurring more frequently on the left side (45%) compared to the right side (38%) [24]. Another study examined the asymmetry in cases of unilateral tongue enlargement [25]. A study was conducted to identify whether there is any impact of velopharyngeal closure due to asymmetry in the length and thickness of the levator veli palatini muscle [26]. The intrinsic velocities of velum raising and lowering movements were analysed for five subjects through cyclic transitions between nasals and fricatives (/s-n-s-n-s-.../). For two male subjects, velum raising was significantly faster than lowering, exhibiting some degree of asymmetry [27]. In contrast, in this study, we focus on ex-

Table 2: Summary of different statistics of number of frames and duration of V1, C, V2 for each of six VCV stimuli

No. of Frames	Stimuli	/ama/	/ana/	/imi/	/ini/	/umu/	/unu/
	Avg (SD)	61.14 (20.6)	62.92 (19.76)	62.32 (21.42)	63.38 (21.3)	62.92 (21.16)	62.44 (21.06)
Max	126	116	120	122	116	116	
Min	26	32	32	32	32	32	
Total	4158	4278	4238	4310	4278	4246	
Duration (in sec) Avg(SD)	V1	0.26 (0.13)	0.27 (0.12)	0.28 (0.13)	0.29 (0.13)	0.26 (0.12)	0.27 (0.12)
	C	0.14 (0.05)	0.14 (0.04)	0.14 (0.05)	0.14 (0.04)	0.14 (0.05)	0.14 (0.05)
	V2	0.3 (0.12)	0.3 (0.11)	0.29 (0.11)	0.29 (0.11)	0.31 (0.12)	0.3 (0.12)

ploring the asymmetry on different aspects of velum dynamics using rtMRI recording of VCV production.

Different modalities, apart from rtMRI, have been used in the past for various studies on velum dynamics, such as X-ray [22], electromyography (EMG) [19], CT scans [8], ultrasound [27] and EMA [20]. There are various pros and cons for using each of these modalities. X-rays and CT scans emit harmful radiation on subjects, making them unsuitable for repeated recordings [28]. Moreover, these techniques are good for recording the data of bones but not ideal for soft tissue and are very difficult to capture real-time data. EMG and EMA require invasive electrodes/sensors to collect the data. Since the velum is soft tissue and inside the vocal tract, it is very challenging and painful for the subject while placing the sensors [29]. Fiberscopy is a highly invasive technique and uncomfortable for subjects [28]. Additionally, it may alter normal speech. Ultrasound might capture real-time and noninvasive data, but it is difficult to effectively capture deep soft tissues like velum. rtMRI serves as an effective modality that offers noninvasive [28, 29, 30, 31], high-resolution, real-time data without obstructing speech, making it suitable to capture real-time 2D MRI video in the mid-sagittal plane to analyse velum movements during the production of nasal sounds within a VCV sequence. Hence, we use rtMRI videos for this study.

This study focuses on asymmetry in velum dynamics using rtMRI video of 68 subjects speaking two nasal sounds, namely /m/ and /n/ in the context of VCV where three vowels, /a/, /i/, and /u/ are used. We propose features from rtMRI images capturing the extent of velar displacement and speed and examine the degree and nature of asymmetry in a symmetric VCV production. Our study reveals that the velar movement is significantly higher (by a factor of  $\sim 2$ ) during velum lowering (V-C transition) compared to that when velum moves up during C-V transition. We also observe that the speed with which velum lowers is significantly higher than that of the speed with which it makes an upward movement in four out of the six VCV stimuli considered in this study.

## 2. Dataset

A subset of the USC 75-speaker speech rtMRI database [32] is used in this study. This subset consists of 68 native American English speakers speaking VCV sequences with C being nasal consonants, namely, /m/ and /n/, and V being the vowels, namely /a/, /i/, and /u/, resulting in a total of six VCV sequences, namely /ama/, /ana/, /imi/, /ini/, /umu/ and /unu/. We use V1 and V2 to indicate the vowel before and after C, respectively. Among 75 speakers available in the database, only 68 (33M+35F) speakers are used in this study, as the velum was not clearly visible in the remaining subjects (sub011, sub018, sub023, sub044, sub049, sub050, sub075). The rtMRI videos are recorded at a frame rate of 83.277 fps, capturing precise

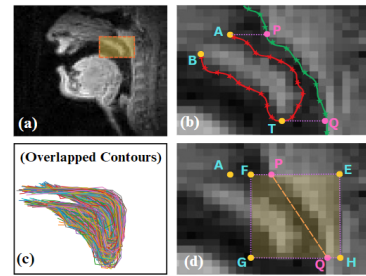


Figure 1: (a) Region of Interest, (b) manual annotation of velum boundary, pharyngeal wall, and velum tip & base points, (c) overlapped velum contours from all six stimuli for an exemplary subject (sub036), (d) illustration of velar feature calculation.

movement of velum with  $84 \times 84$  pixel resolution. Each pixel has  $2.4 \times 2.4 \text{ mm}^2$  as spatial area. To extract relevant rtMRI image frames from a VCV recording, the beginning and end timestamps of the V1, C, V2 segments were identified for each subject across all six VCV sequences. This segmentation was done with the help of Praat<sup>1</sup> by observing the waveform and spectrogram of the spoken utterance. The duration of V1, C, V2 along with different statistics of number of frames in these segments are summarised in Table 2. From the extracted frames, only even-numbered frames (i.e., alternate frames) were selected for manual annotation using MATLAB Graphical User Interface (GUI)[33]. Thus, a total of 2079, 2139, 2119, 2155, 2139, and 2123 frames are used for manual annotation in this study for /ama/, /ana/, /imi/, /ini/, /umu/ and /unu/, respectively. The region of interest (RoI) as shown in Figure 1(a) are used for manual annotation. A contour marking the boundary between the velum and the air as well as a contour representing the pharyngeal wall are manually marked (shown in red and green contours in Figure 1(b) respectively) along with two basepoints of the velum (A and B) and velum tip (T), indicated by yellow colored dots in the same figure. From the velum boundary annotation, we found that the velum exhibits different shapes, as reported in [11]. For instance, during /a/ production, leaf, crook, straight, butt and rat trail shapes are observed for 41, 8, 9, 6, 4 speakers, respectively. The shape of the velum varies depending on the subject and the stimuli used.

## 3. Study details

In order to represent velum movement during VCV production, a feature related to the image intensity in a ROI centred around the velum (as shown in Figure 1) is calculated. Given a sequence of  $M$  rtMRI images from a video recorded during a VCV production, let  $f[m]$ ,  $1 \leq m \leq M$  denotes the velar feature value calculated from the  $m$ -th rtMRI image (an illustrative plot of  $f[m]$  vs  $m$  is shown in Figure 2). Below we describe the steps for calculating  $f[m]$ .

Let  $\mathcal{I}_m[k, l]$ ,  $1 \leq k, l \leq 84$  denotes the pixel value of the  $m$ -th rtMRI image at the pixel location  $(k, l)$ . From the manually annotated velum end-point (A) near the velum base, a horizontal straight line is drawn and the point (P) is identified where this line intersects the pharyngeal wall contour (as shown Figure 1(b)). Similarly, point Q is identified starting from the manually marked velum tip point T. Now, using points, A, P, and Q, a quadrilateral, named EFGH (as shown Figure 1(d)) is constructed in the following manner. At first, the mid-point (F) between points A and P is identified. Then point G is identi-

<sup>1</sup><https://www.fon.hum.uva.nl/praat>

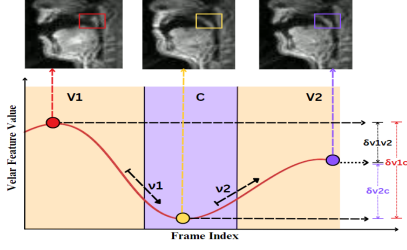


Figure 2: Illustration of the velar feature calculated from velum dynamics in this study.

fied as the crossover point of the vertical line going through F and the horizontal line passing through point Q. Point E on the line joining F and P is identified to ensure that the distance between G and Q is identical to the distance between P and E. Finally, point H is identified as the intersection point of the vertical line passing through point E and horizontal line passing through point Q. Once, the quadrilateral EFGH is determined, the  $f[m]$  is calculated as follows:

$$f[m] = \sum_{(k,l) \in EFGH} \mathcal{I}_m[k, l]$$

For calculation of  $f[m]$ , we do not include part of the velum lying in the region to the left of the line FG, although velum base point is A (near the base of the velum) is to the left of FG. This is because, through intensity based velar feature, we would like to capture the motion of the velum. However, near the base of the velum, the velum movement is minimal (as shown in Figure 1(c) for all six VCV sequences together for an exemplary subject) and most of the intensity change due to velum movement is contributed by the pixels lying to the right side of the line FG. Pixel intensity based feature has been used in the past for representing velum movement. While some of them have used triangular area of interest [16], some other have used rectangular area of interest [34] as well similar to the one proposed in our study.

In order to examine the asymmetry, we investigate how the velar feature  $f[m]$  changes from V1 to C and back to V2. This is done using the extent and speed with which  $f[m]$  changes from V1 to C, and back to V2.

### 3.1. Extent of change in $f[m]$

The extent of velar feature change is calculated by first computing statistics of  $f[m]$  in each of the three (V1, C, V2) segments, denoted by  $\xi_{V1}$ ,  $\xi_C$ , and  $\xi_{V2}$ , respectively. The extent of velar feature change is then calculated using  $\delta_{V1C} = \xi_{V1} - \xi_C$ ,  $\delta_{V2C} = \xi_{V2} - \xi_C$ . These are illustrated in Figure 2 using the velar feature sequence shown there. We carry out statistical test to examine if the mean values of  $\delta_{V1C}$  and  $\delta_{V2C}$  are individually nonzero. This is done using t-test with null hypothesis being the mean value is equal to zero. We also carry out statistical test to examine if  $\delta_{V1C}$  and  $\delta_{V2C}$  are statistically significantly different between themselves (also show in Figure 2), i.e.,  $\delta_{V1V2} = \delta_{V1C} - \delta_{V2C} \neq 0$ . This is done using t-test with null hypothesis being that the mean value of  $\delta_{V1V2}$  is equal to zero. In order to calculate  $\xi_{V1}$ ,  $\xi_C$ , and  $\xi_{V2}$ , two different types of statistics are computed from  $f[m]$  as described below

**Extreme Value (EV):** In this case, we compute the maximum value of  $f[m]$  in V1 and V2 segments, while we compute

minimum value of  $f[m]$  in C segment (shown in Figure 2), i.e.,

$$\xi_{V1} = \max_{m \in V1} f[m], \quad \xi_{V2} = \max_{m \in V2} f[m], \quad \xi_C = \min_{m \in C} f[m].$$

Velum lowers during nasal production causing the intensity value to decrease in EFGH quadrilateral compared those during vowel production before and after nasal.

**Average Value (AV):** In this case, we compute the average of  $f[m]$  in each of the three segments separately. Thus, using AV, we obtain

$$\xi_{V1} = \sum_{m \in V1} \frac{f[m]}{N_{V1}}, \quad \xi_{V2} = \sum_{m \in V2} \frac{f[m]}{N_{V2}}, \quad \xi_C = \sum_{m \in C} \frac{f[m]}{N_C},$$

where  $N_{V1}$ ,  $N_C$ ,  $N_{V2}$  are the number of rtMRI images in V1, C, V2, respectively.

In the case of AV, an average position of the velum during each of the V1, C, V2 segments is captured, while in the case of EV,  $\xi_{V1}$  and  $\xi_{V2}$  is used to capture the position of velum when nasal passage is closed during vowel production and  $\xi_C$  is used to capture the position when velum lowers completely during nasal production opening, the nasal passage.

### 3.2. Speed of change in $f[m]$

While extent of change in  $f[m]$  captures different position of the velum during VCV production, the speed with which velum moves from opening to closing position is not represented in the  $\xi_{V1}$ ,  $\xi_C$ , and  $\xi_{V2}$ . The speed of velar feature change is calculated by computing the magnitude of velocities [35],  $\nu_1$  and  $\nu_2$  at the V1-C and C-V2 transitions separately. This is done using  $L$  velar feature values on either side of the transition point. Suppose  $k_1$  and  $k_2$  are the two frame indices indicating the V1-C and C-V2 transitions, then

$$\nu_1 = \left| \frac{\sum_{l=1}^L l \times (f[k_1 + l] - f[k_1 - l])}{\sum_{l=1}^L l^2} \right|.$$

Similarly,  $\nu_2$  is also calculated. In order to examine the asymmetry in terms of speed of velar feature change, we carry out statistical test to examine if  $\nu_1$  is statistically significantly different from  $\nu_2$  using a paired t-test with the null hypothesis that the mean value of their difference ( $\nu_1 - \nu_2$ ) is equal to zero.

## 4. Results and Discussion

### 4.1. Asymmetry in the extent of change in velar feature using EV statistics

The normalized histograms of  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  calculated using EV statistics of velar feature (as explained in Section 3) are shown in Figure 3, using blue, green, and red colors, respectively, for each of the six VCV stimuli. A kernel density estimate of the probability density function (PDF) is also shown in the same color (filled density graphs). A vertical dashed line at zero point on the x-axis is shown to indicate how much the histograms of  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  lie on the positive part of the x-axis. It is clear from Figure 3 that the  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  are majorly positive valued. In particular, 97.06%, 98.53%, 100%, 97.06%, 98.53%, and 97.06%, of  $\delta_{V1C}$  are positive valued for /ama/, /imi/, /umu/, /ana/, /ini/, and /unu/, respectively. These percentages are 91.18%, 89.71%, 91.18%, 98.53%, 89.71%, and 83.82% for  $\delta_{V2C}$  and 75.00%, 73.53%, 80.88%, 76.47%, 80.88%, and 79.41% for  $\delta_{V1V2}$ . The average (Avg) and standard deviation (SD) are also shown for each of

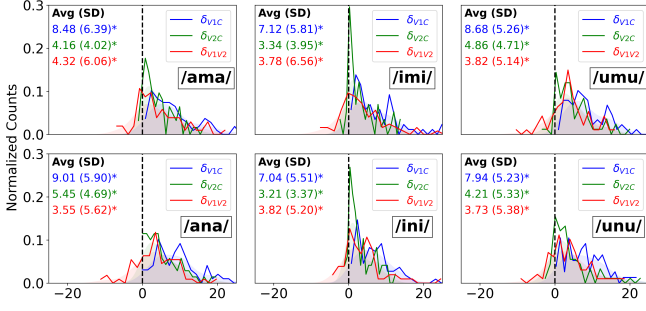


Figure 3: Normalized histograms of  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  using EV based statistics of the intensity feature

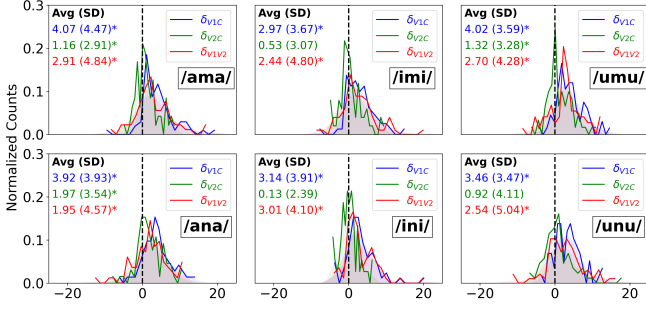


Figure 4: Normalized histograms of  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  using AV based statistics of the velar feature

them in the every sub-figure in Figure 3. A star next to the SD value indicates that the corresponding average value is significantly ( $p < 0.01$ ) greater than zero. It is interesting that  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  for each of the six VCV stimuli have average value which is significantly greater than zero. This suggests that there is a significant drop from  $\xi_{V1}$  to  $\xi_C$  suggesting the lowering of velum from V1 segment to the nasal segment. Similarly, a significant drop from  $\xi_{V2}$  to  $\xi_C$  indicates the upward movement of the velum from nasal segment to V2 segment. However, it can be observed that average value of  $\delta_{V1C}$  is larger than that of  $\delta_{V2C}$  by a factor of 2.03, 2.13, 1.78, 1.65, 2.19, and 1.88 for /ama/, /imi/, /umu/, /ana/, /ini/, and /unu/, respectively. This, in turn, suggests that for the production of vowel after the nasal segment, the velum does move upward but only up to approximately half way of where it was during the production of vowel before the nasal segment. This is true for all of the six VCV stimuli, which is also clear from the significantly ( $p < 0.01$ ) positive value of  $\delta_{V1V2}$ , as shown by red color in Figure 3, indicating the asymmetry in downward and upward movement of the velum during VCV production.

#### 4.2. Asymmetry in the extent of change in velar feature using AV statistics

Similar to Figure 3, the normalized histograms of  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  calculated using AV are shown in Figure 4. We observed a similar pattern in the Avg and SD values as shown in blue, green and red colored text in Figure 4. The histograms lie primarily on the positive side of the x-axis as reflected in the positive average values. However, unlike EV in Figure 3, the average value of  $\delta_{V2C}$  is not significantly greater than zero in the case of /imi/, /ini/, and /unu/. This suggests that post nasal segment, velum does not move upward significantly. However, for rest of the cases, average value of  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  are significantly greater than zero. In particularly, considering significantly positive average value of  $\delta_{V1V2}$ , it can be con-

Table 3: Avg(SD) values of  $\nu_1 - \nu_2$  and percentage +ve and -ve values of  $\nu_1 - \nu_2$  for each VCV stimulus for different  $L$ .

$L$	$\nu_1 - \nu_2$	/ama/	/ana/	/imi/	/ini/	/umu/	/unu/
$L = 3$	Avg (SD)	0.19 (0.59)	0.02 (0.64)	0.13 (0.52)	0.21 (0.47)	0.18 (0.52)	0.07 (0.56)
	(+) / (-)	60.3/39.7	48.5/51.5	54.4/45.6	63.2/36.8	63.2/36.8	58.8/41.2
	Avg (SD)	0.20 (0.56)	0.00 (0.52)	0.13 (0.48)	0.21 (0.41)	0.20 (0.49)	0.08 (0.48)
$L = 4$	(+) / (-)	64.7/35.3	50/50	58.8/41.2	67.6/32.4	66.2/33.8	57.4/42.6
	Avg (SD)	0.20 (0.54)	0.01 (0.48)	0.15 (0.44)	0.20 (0.38)	0.19 (0.48)	0.09 (0.47)
	(+) / (-)	63.2/36.8	50/50	61.8/38.2	73.5/26.5	66.2/33.8	60.3/39.7
$L = 5$	Avg (SD)	0.18 (0.49)	0.01 (0.45)	0.14 (0.39)	0.19 (0.34)	0.17 (0.46)	0.10 (0.45)
	(+) / (-)	63.2/36.8	51.5/48.5	65.2/34.8	71.2/27.9	64.7/35.3	58.8/41.2

cluded, just like in the case of EV, that the amount of velum movement is not symmetric during its downward and upward motion in VCV production. It is important to note that, due to averaging unlike extreme value, the average values of  $\delta_{V1C}$ ,  $\delta_{V2C}$ , and  $\delta_{V1V2}$  using AV (Figure 4) are less than those using EV (Figure 3).

#### 4.3. Asymmetry in the speed of change in velar feature

Four different values of  $L$  are used for the study of asymmetry in the speed of change in velar feature, namely,  $L = 3, 4, 5, 6$ . For each value of  $L$ ,  $\nu_1 - \nu_2$  (as defined in section 3.2) is calculated for each VCV stimulus for every speaker. The Avg and SD values of  $\nu_1 - \nu_2$  are shown in Table 3 along with the percentages of positive and negative values, indicating  $\nu_1 > \nu_2$  and  $\nu_1 < \nu_2$ , respectively. It is found that except for /ana/ and /unu/, the average value of  $\nu_1 - \nu_2$  is significantly ( $p < 0.01$ ) greater than zero for all values of  $L$  other than the case of /ama/ for  $L = 3$  and /imi/ for  $L = 3, 4$ . It is also evident from the percentages of +ve and -ve values. This suggests that the speed with which velum lowers is significantly higher than that with which it makes an upward movement post nasal segment suggesting an asymmetry in the velar movement speed in VCV production. In particular, for  $L=5$ , the average  $\nu_1$  is 1.58, 1.0, 1.45, 1.78, 1.5, and 1.23 times higher than average  $\nu_2$  for /ama/, /ana/, /imi/, /ini/, /umu/, and /unu/, respectively.

## 5. Conclusions

Using rtMRI video during VCV production from 68 speakers, this study reveals that the velar displacement and velocity is not symmetric during velum lowering (V-C transition) and velum raising (C-V transition). A significant higher velum lowering happens during V-C transition with relatively higher speed than those during C-V transition. This suggests a stronger carryover effect as vowel post nasal segment remains more nasalized compared to the vowel before nasal segment. Although results from this study does not match with that report by [27] (probably due to different tasks and less number of subjects used in that study), they are similar to the findings reported by Rossato et al. [20]. While this study only deals with displacement and velocity of velar dynamics, higher-order dynamical characteristics remain to be investigated. Further study needs to be carried out on a larger pool of subjects to examine whether there is any significant change in velum shape and size during VCV production and if any asymmetry is observed in such velar features. The sensitivity of the findings in this study to the annotation error and the anatomical variability needs to be investigated as well. These are the parts of our future work.

## 6. References

- [1] M. G. Busà *et al.*, “Coarticulatory Nasalization and Phonological Developments: Data from Italian and English Nasal-Fricative Sequences,” in *Experimental Approaches to Phonology*. Oxford University Press, 2007, pp. 155–174.
- [2] —, “Vowel Nasalization and Nasal Loss in Italian,” in *Proceedings of the 15th International Congress of Phonetic Sciences*, 2003, pp. 711–714.
- [3] J. M. Inouye, S. Dabbaghchian, J. L. Perry, D. P. Kuehn, J. B. Moon, and B. P. Sutton, “A Computational Model Quantifies the Effect of Anatomical Variability on Velopharyngeal Function,” *Journal of Speech, Language, and Hearing Research*, vol. 58, no. 4, pp. 1119–1133, 2015.
- [4] A. Gauster, Y. Yunusova, and D. Zajac, “The Effect of Speaking Rate on Velopharyngeal Function in Healthy Speakers,” *Clinical Linguistics & Phonetics*, vol. 24, no. 7, pp. 576–588, 2010.
- [5] X. Niu, A. Kain, and J. P. H. van Santen, “Estimation of the Acoustic Properties of the Nasal Tract during the Production of Nasalized Vowels,” in *Interspeech*, 2005, pp. 1045–1048.
- [6] L.-S. Su, K.-P. Li, and K.-S. Fu, “Identification of Speakers by Use of Nasal Coarticulation,” *The Journal of the Acoustical Society of America*, vol. 56, no. 6, pp. 1876–1882, 1974.
- [7] A. J. S. Teixeira and F. A. C. Vaz, “European Portuguese Nasal Vowels: An EMMA Study,” in *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech)*, 2001, pp. 1483–1486.
- [8] T. Vampola, J. Horáček, V. Radolf, J. G. Švec, and A.-M. Laukkanen, “Influence of Nasal Cavities on Voice Quality: Computer Simulations and Experiments,” *The Journal of the Acoustical Society of America*, vol. 148, no. 5, pp. 3218–3231, 2020.
- [9] H. M. Stoakes, J. M. Fletcher, and A. R. Butcher, “Nasal Coarticulation in Biniw Kunwok: An Aerodynamic Analysis,” *Journal of the International Phonetic Association*, vol. 50, no. 3, pp. 305–332, 2020.
- [10] G. Zellou and M. Tamminga, “Nasal Coarticulation Changes over Time in Philadelphia English,” *Journal of Phonetics*, vol. 47, pp. 18–35, 2014.
- [11] K. J. Kotlarek, A. E. Haessler, K. E. Hildebrand, and J. L. Perry, “Morphological Variation of the Velum in Children and Adults Using Magnetic Resonance Imaging,” *Imaging Science in Dentistry*, vol. 49, no. 2, pp. 153–158, 2019.
- [12] I. Korhan, S. Gode, R. Midilli, and O. K. Basoglu, “The Influence of the Lateral Pharyngeal Wall Anatomy on Snoring and Sleep Apnoea,” *Journal of the Pakistan Medical Association*, vol. 65, no. 2, pp. 125–130, 2015.
- [13] J. Dang and K. Honda, “An Improved Vocal Tract Model of Vowel Production Implementing Piriform Resonance and Transvelar Nasal Coupling,” in *Proceedings of the Fourth International Conference on Spoken Language Processing (ICSLP)*, vol. 2. IEEE, 1996, pp. 965–968.
- [14] S. Liao, P. Hoole, and J. Harrington, “Comparison of Velum Control in /an/-rime Words between Chengdu Dialect and Standard Mandarin,” in *Proceedings of the 13th International Seminar on Speech Production (ISSP)*, 2024.
- [15] A. Purnomo, N. Ebbutt, C. Purnomo, J. Islam, G. de Boer, and B. Gick, “Comparing Velum Velocity in Québécois French Nasals,” *Canadian Acoustics*, vol. 50, no. 3, pp. 98–99, 2022.
- [16] M. Oh, D. Byrd, and S. S. Narayanan, “Leveraging Real-Time MRI for Illuminating Linguistic Velum Action,” in *Proceedings of Interspeech*, 2021, pp. 3964–3968.
- [17] J. Dang, J. Wei, K. Honda, and T. Nakai, “A study on transvelar coupling for non-nasalized sounds,” *The Journal of the Acoustical Society of America*, vol. 139, no. 1, pp. 441–454, 2016.
- [18] M. S. Barlaz, L. Goldstein, and S. S. Narayanan, “The Emergence of Nasal Velar Codas in Brazilian Portuguese: An rt-MRI Study,” in *Proceedings of Interspeech*, 2015, pp. 2660–2664.
- [19] J. Freitas, A. Teixeira, S. Silva, C. Oliveira, and M. S. Dias, “Velum Movement Detection Based on Surface Electromyography for Speech Interface,” in *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS)*. SciTePress, 2014, pp. 13–20.
- [20] S. Rossato, P. Badin, and F. Bouaouni, “Velar Movements in French: An Articulatory and Acoustical Analysis of Coarticulation,” in *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, 2003, pp. 3141–3144.
- [21] D. P. Kuehn and J. B. Moon, “Velopharyngeal closure force and levator veli palatini activation levels in varying phonetic contexts,” *Journal of Speech, Language, and Hearing Research*, vol. 41, no. 1, pp. 51–62, 1998.
- [22] J. Dang and K. Honda, “INVESTIGATION OF THE ACOUSTIC CHARACTERISTICS OF THE VELUM FOR VOWELS,” in *Proceedings of the International Conference on Spoken Language Processing (ICSLP)*. International Speech Communication Association, 1994, pp. 603–606.
- [23] N. Miller, C. Reyes-Aldasoro, and J. Verhoeven, “Lateral Asymmetry in the Articulation of British English Speech Sounds: An Electropalatographic Study,” *Journal of the International Phonetic Association*, pp. 1–19, 2025.
- [24] J. Verhoeven, P. Mariën, I. De Clerck, L. Daems, C. C. Reyes-Aldasoro, and N. Miller, “Asymmetries in Speech Articulation as Reflected on Palatograms: A Meta-Study,” in *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)*, 2019, pp. 2820–2824.
- [25] K. H. Ausermann and E. Machtens, “The Influence of Tongue Asymmetries on the Development of Jaws and the Position of Teeth,” *International Journal of Oral Surgery*, vol. 3, no. 5, pp. 261–265, 1974.
- [26] N. Tahmasebifard, J. L. Perry, D. P. Kuehn, X. Fang, and B. P. Sutton, “Evaluation of the Symmetry of the Levator Veli Palatini Muscle and Velopharyngeal Closure Among a Nonleft Adult Population,” *The Cleft Palate-Craniofacial Journal*, vol. 58, no. 6, pp. 728–735, 2021.
- [27] P. Birkholz and C. Kleiner, “Velocity Differences between Velum Raising and Lowering Movements,” in *Speech and Computer: 23rd International Conference, SPECOM 2021, St. Petersburg, Russia, September 27–30, Proceedings*, ser. Lecture Notes in Computer Science, vol. 12997. Springer, 2021, pp. 70–80.
- [28] S. G. Lingala, B. P. Sutton, M. E. Miquel, and K. S. Nayak, “Recommendations for Real-Time Speech MRI,” *Journal of Magnetic Resonance Imaging*, vol. 43, no. 1, pp. 28–44, 2016.
- [29] J. Freitas, A. Teixeira, S. Silva, C. Oliveira, and M. S. Dias, “Detecting Nasal Vowels in Speech Interfaces Based on Surface Electromyography,” *PLOS ONE*, vol. 10, no. 6, p. e0127040, 2015.
- [30] J. L. Perry, B. P. Sutton, D. P. Kuehn, and J. K. Gamage, “Using MRI for Assessing Velopharyngeal Structures and Function,” *The Cleft Palate-Craniofacial Journal*, vol. 51, no. 4, pp. 476–486, 2014.
- [31] J. L. Perry, K. Mason, B. P. Sutton, and D. P. Kuehn, “Can Dynamic MRI Be Used to Accurately Identify Velopharyngeal Closure Patterns?” *The Cleft Palate-Craniofacial Journal*, vol. 55, no. 4, pp. 499–507, 2018.
- [32] Y. Lim, A. Toutios, Y. Bliesener, Y. Tian, S. G. Lingala, C. Vaz, T. Sorensen, M. Oh, S. Harper, W. Chen *et al.*, “A multispeaker dataset of raw and reconstructed speech production real-time mri video and 3d volumetric images,” *Scientific data*, vol. 8, no. 1, p. 187, 2021.
- [33] A. K. Pattem, A. Illa, A. Afshan, and P. K. Ghosh, “Optimal sensor placement in electromagnetic articulography recording for speech production study,” *Computer speech & language*, vol. 47, pp. 157–174, 2018.
- [34] N. Wong, M. Fu, Z.-P. Liang, R. Shosted, and B. P. Sutton, “Observations of perseverative coarticulation in lateral approximants using mri.” in *INTERSPEECH*, 2013, pp. 612–616.
- [35] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey *et al.*, “The htk book,” 1999.