Articulatory Setting of the Tongue for Eastern Versus Western Japanese

Atsushi Ueno   s1160031

Abstract

The underlying setting of the articulators in speech is called the articulatory setting. This setting varies across languages and dialects. This research focuses on Japanese articulatory setting of the tongue, especially the differences between eastern and western Japanese. The method of this research is mainly collecting ultrasound data of subjects, and analyzing the tongue’s center of gravity using software such as Edgetrak and MATLAB. Edgetrak is used for getting the coordinates of the tongue surface and MATLAB is used for calculating the center of gravity. The results of our analysis show that the distance from the center of gravity of the tongue to the top of the ultrasound probe is greater, on average, for western Japanese subjects than for eastern Japanese subjects. In other words, the rest position of the tongue is higher for western Japanese than it is for eastern Japanese. In addition, we found an effect of the preceding vowel on the rest position of the tongue.

1 Introduction

Even within the same language, there are different features of speaking (intonation, articulation, etc.) depending on the dialect. Of course, Japanese is no exception, and the differences between those features are clear between eastern and western Japanese dialects. For example, in many eastern Japanese dialects, the pitch gradually declines toward the end of a declarative sentence. On the other hand, speakers of some western Japanese dialects tend to raise the pitch of the word at the end of a sentence. Thus, there are such distinctions between eastern and western Japanese. So, the influence of such distinctions on each articulator needs to be measured. It is hypothesized that these different ways of speaking (i.e., different dialects) have different articulatory settings. If speakers are aware of how the settings differ across languages and dialects, it may help them when they are acquiring a new language.

With the improvement of measurement techniques such as ultrasound, researchers have been able to study and measure the articulatory setting of various languages. According to a study of the articulatory setting of French and English monolinguals and bilinguals [1], tongue tip height of English speakers was higher than French speakers during speech rest position (between utterances). Furthermore, English speakers tended to protrude upper and lower lips more than French speakers during rest position. As just described, it is apparent that there are differences between articulatory settings across languages.

One of the key things that cannot be overlooked is how we define the underlying tongue position (i.e., the articulatory setting). The research noted above [1] focused on the tongue tip height from the ultrasound probe during rest position between utterances as the method of finding out articulatory setting of the tongue. One problem with that method is the difficulty in finding the alveolar ridge in the ultrasound images. If one cannot find consistent bone points in the ultrasound image, it is difficult to consistently measure the same point on the tongue. So, in this research, we considered the center of gravity as the definition of tongue position. The reason why we use center of gravity of the tongue is because it is easier to see the movement of the whole tongue and how different tongue rest positions are affected by neighboring vowels.

Considering western Japanese, especially the people living in Kansai, one feature of speaking is that assimilation in diphthongs, such as “sugee” for “sugoi” or “samii” for “samui” hardly occurs. On the other hand, eastern Japanese, especially the people living in Tokyo, often use assimilation of diphthongs like the above [2]. In general, it is known that the tongue position when Japanese pronounce long vowels such as “i” and “e” becomes more forward [3]. Consequently, it is hypothesized that eastern Japanese tongue rest position and its center of gravity would be slightly more forward than western Japanese because they use long vowel frequently such as “i” and “e”.

Thus, the topics we cover in this research are as follows: 1) Finding the difference between articulatory setting of the tongue for eastern and western Japanese. 2) Finding the effect of each preceding vowel on the tongue’s rest position.
2 Method

2.1 Participants

We collected data from nine western Japanese subjects. We used the eastern Japanese data that had been collected for earlier research [4]. It is difficult to define the border of eastern and western Japan because there are many methods of how to divide up Japan depending on the viewpoint (e.g. culture, food and language). In principle, we need to decide that border from a linguistic point of view. Even that, though, is very difficult to do because there are so many dialects of Japanese. However, in this case, we defined Tohoku and Kanto as eastern Japan and Chubu, Kinki, Shikoku, Chugoku and Kyushu as western Japan. It is possible that one or two eastern subjects spoke with a western dialect, but not likely. All the western Japanese participants were undergraduate students and have the dialects of each rural area.

2.2 Stimuli

The stimuli that we used for data collection were short Japanese sentences identical to those used in the eastern Japanese data collection [4]. We did three trials of 30 Japanese sentences each; each trial had the same set of 30 sentences rearranged in random order. One example of a stimulus sentence is (Mr.Chiba is tone deaf). Because we suspected that articulatory setting is affected by preceding and following consonants and vowels, we controlled for the sound at the beginning and end of sentences (to balance the numbers of each consonant and vowel). So, we used six PowerPoint slides for each sentence-final vowel, i.e. あ [a], い [i], う [u], え [e], お [o], to balance the numbers of vowels. Stimuli were displayed on a Mac laptop and presented to subjects in automatic PowerPoint presentation, one sentence at a time.

2.3 Apparatus

Ultrasound (Toshiba Famio 8) was used to record the movement of the tongue in speech. Two video cameras (JVC GZ-HD7 from the front and Panasonic HDC-TM750 from the side) were used to record the head movement of each subject. The front image of the face was collected for future analysis of lip movements. Two Mac computers (iMac and MacBook Pro) were used, the latter for displaying stimuli, and the former for recording the combination of ultrasound movie and audio. A microphone (Shure SM10A) was used to record the subject’s utterances. An audio and visual mixer (Canopus ADVC-700) combined the ultrasound movie and the audio of each subject. A specially made helmet and glasses that I constructed were used to track four pink dots that were attached to those and we could determine the angle between the line of the upper two dots and that of the lower two dots. Also, four lip markers were used to measure the subject’s mouth aperture. Two halogen lights (LPL VL-1300+PRO4) were used to shine light on the subject indirectly using light reflectors.

Figure 1: Helmet holding ultrasound probe and glasses used for data collection

2.4 Procedure

2.4.1 Data Collection

We collected western Japanese data to compare with previously collected eastern Japanese data [4]. First, we explained to the subjects that what they should do is simply read some sentences, and we hid the purpose of this research from the subject (so as not to influence their usual speaking pattern). Second, the subject had to put on some equipment (e.g., helmet, lip markers, etc.), as discussed above. Third, we checked whether all equipment was working normally or not. Before displaying stimuli, the subject had to swallow so that we had a record of where the palate is (for future research). Finally, we struck a clapper and displayed the stimuli.

2.4.2 Data Analysis

Rest position frame For western Japanese, first of all, we had to look for the rest position frame of the tongue from the ultrasound movie. The definition of rest position was the condition in which the motion of the tongue stopped for more than 0.3 seconds (10 frames) between each pair of sentences. The number
of possible rest position frames for each subject was 30 because one trial that we used for stimuli had 30 sentences. However, there were only from 5 to 15 usable rest position frames per subject because of swallowing, etc. With 3 trials (90 sentences) per subject, we could collect about 30 suitable rest position frames per subject. For eastern Japanese, the same method [4] was used.

**EdgeTrak** To calculate the center of gravity of the tongue, we first needed to find the edge of the tongue’s surface in the ultrasound image and get the coordinates of that edge. We could see the boundary of a white and black line in ultrasound image. This black line shows the actual tongue surface. Tracking this black line for each rest position frame, we could get the coordinates of the tongue’s surface at rest between sentences.

**MATLAB calculations** We used MATLAB and wrote code for calculating the center of gravity. The process of calculating the center of gravity is as follows: 1) Importing the coordinates of the tongue’s surface from EdgeTrak, 2) Specifying the Region Of Interest (ROI), 3) Calculating the center of gravity of that ROI. In this research, we specified the ROI as bounded by 4 lines, the tongue’s surface, the left and right straight line of the ultrasound’s field of view and the lower parabola showing the surface of the ultrasound probe. Because of the sublingual air space (air under the tongue tip), it is often difficult to see the tongue tip clearly with ultrasound. In specifying the ROI, we ignored the sublingual air space and included it as part of the tongue. For this reason, all center of gravity measurements have some error associated with them, but this error is not systematic (i.e., it exists for all frames and should not affect the comparison results).

**Correcting error for western Japanese** The recording method that we used in this experiment had the possibility of probe rotation from its initial position. If that were to happen, we would not be calculating the true center of gravity of the tongue. To tackle this problem, we used the following method. Each subject was recorded from the side using a video camera. The reason for recording was to know the angle between two blue bars: one firmly attached to the glasses (hence measuring head rotation) and the other blue bar attached to wood that held the ultrasound probe (hence measuring probe rotation) - see Figure 1. The wood holding the probe was attached to a helmet that the subjects wore. By knowing the angle between the blue bars, we could calculate how much the probe rotated relative to the head in each rest position frame, and then we could correct for this movement.

The process of correcting error for western Japanese was as follows: 1) Find the rest position side-video frame at the same time as the ultrasound tongue rest position frame. 2) Calculate central coordinates of each pink dot. 3) Calculate the angle between the line defined by the upper two dots and that defined by the lower two dots. 4) Correct the error and rotate the center of gravity. For finding the central coordinates of each pink dot in step 2, a C language program was used [5]. As for steps 3 and 4, we used MATLAB to make those calculations.

**Correcting error for eastern Japanese** As for eastern Japanese, we used a somewhat different method. That was because the head motion had been tracked using a Vicon MX point-tracking system instead of a video camera. Therefore, we used Vicon coordinates of the glasses, chin and probe. The Vicon data had 3D coordinates. However, we used only the y-z plane because we were interested in rotation in the midsagittal plane (i.e., the same type of coordinates as the side movie). The process of correcting error for eastern Japanese was as follows: 1) Find the coordinates from a CSV file for the markers on the glasses, chin and probe. The Vicon data had 3D coordinates. However, we used only the y-z plane because we were interested in rotation in the midsagittal plane. 2) Calculate the angle between the extended line of two probe coordinates and the extended line of two glass coordinates. 3) Correct the error and rotate the center of gravity.

**Rotating tongue’s center of gravity** Based on the angle data as discussed above, we rotated the coordinates of the center of gravity based on the angle in the first rest position frame.

**Calculating average and normalization** Finally, we calculated the average of each subject’s center of gravity coordinates for rest positions sorted by preceding vowel, i.e. は [a], い [i], う [u], え [e], そ [o]. In addition, we calculated the distance between top of probe to center of gravity (see Figure 2). As for this distance, we had to normalize because each subject’s face size is not the same. In general, vocal tract size is correlated with face length. So, if we set the standard of face size, we can normalize easily.

The process of western Japanese normalization was as follows: 1) Find a clear image of each subject from the movie. 2) Find the center point between the left
and right eye, and calculate the distance between that center point and the chin. 3) Set the basis value equal to one, i.e., set the distance of the subject who has the longest face equal to one. 4) Calculate the relative value from the basis value, i.e., divide other subject’s distance by the distance of the subject who has the longest face. 5) Divide the distance between the top of the probe and the center of gravity by that relative value.

As for eastern Japanese, the method of calculating the average was the same, but normalization was slightly different from western Japanese. The only difference was how to find the center point. We found the center point by using left side and right side glasses markers because we did not have video images of eastern Japanese subjects. It was same method after step 2.

3 Results

We show the following results: 1) the average distance between the tongue’s center of gravity to top of the probe, 2) the actual coordinates of the center of gravity, 3) X and Y coordinate difference between the center of gravity and the top of the probe.

The first results can be seen in Figure 3. As for this result, we calculated the average for eastern and western Japanese speakers. According to this result, the western Japanese distance is approximately 7 mm longer than eastern Japanese.

The second results can be seen in Figure 4. This result is shown by scatter plot. It shows the average location of the tongue’s center of gravity for each subject. In addition, two big points show the average of eastern and western Japanese. Seeing the coordinate locations of the center of gravity, the difference between eastern and western is clear.

The third result can be seen in Figures 5 and 6. Figure 5 focused on the X coordinates, and Figure 6 focused on the Y coordinates. Looking at Figure 5, we can see the difference of length across the vowels. For example, the rest position of the tongue after vowels “う [u]” and “お [o]” is farther back than the position after the other three vowels. As for Figure 6, however, all of the bar heights do not change so much.

4 Discussion

First of all, the complicated nature of this experimental research caused us to greatly reflect on our methodology. We considered that the experiment needed standardization of experimental methodology across all subjects. One difficulty that we had is that the number of subjects was small for both eastern and western Japanese, making it difficult to generalize the results to larger populations. In hindsight, we should have prepared more western Japanese subjects in expectation of obtaining data that could not be used. In addition, we should have used the same head tracking method for eastern Japanese as was used for western Japanese so that normalization of subjects would be more straightforward. Furthermore, the geographical area was too large for simple labels such as eastern and western Japanese. It is not correct to say that all eastern and western Japanese speak one dialect respectively. Therefore, there may exist various articulatory settings among eastern and western Japan respectively. By narrowing the geographical area to, for example,
Figure 3: Distance between tongue’s center of gravity and top of probe. The right bar is eastern Japanese and the left bar is western Japanese.

Figure 4: A plot of the coordinates of tongue’s center of gravity. Small points show each subject’s average, and the two big points show eastern and western Japanese average. The units are in pixels where 4 pixels = 1 mm.

Figure 5: The average difference of X coordinate distance between center of gravity and top of probe. It shows all vowel average at the top, and rest position after each vowel below that.

Figure 6: The average difference of Y coordinate distance between center of gravity and top of probe. The graph order is the same as Figure 5 but rotated 90°.
Tokyo and Osaka, we could eliminate the difference of articulatory setting within our groups and would focus research on articulatory setting of limited areas.

We can say the data collection of western Japanese was accurate because the change of the angle of the probe relative to the head was almost nothing and so we did not need to rotate the images. The maximum difference of the angle in all western data was about 0.9°, and the difference of X coordinates before and after rotating was about 0.4 mm. Therefore, rotation could be ignored. This result proved the data collection system using the helmet and glasses accurately maintains the probe angle relative to the head.

Another challenge we encountered when doing this experiment was that we needed to normalize the data, but because both normalization method were not exactly the same, we had a doubt about the accuracy. But because eastern and western Japanese average head size are not different, we solved that problem by calculating the average over all subjects. For calculating average, some female data became an obstacle because there is a distinction of head size between male and female. Thus, data from two female subjects of eastern Japanese were ignored. In addition, three western Japanese subjects’ data were not available because of defects in the experiment (e.g. pilot data, difficult to see the tongue surface in Edgetrak, etc.)

We expected that each vowel’s tongue position affects the rest position that follows. For example, the rest position after pronouncing “∫ [ur]” or “Ę [o]” would be more back as compared to other vowels because those two are back vowels. As expected, vowel tongue position does affect rest position as shown by Figure 5. In addition, eastern Japanese seem to be more susceptible to the vowel tongue position because the change of X coordinates by vowel is bigger than that for western Japanese. However, the Y coordinates did not change so much, as shown by Figure 6. It means that rest position is more susceptible to forward and back movement.

Contrary to our expectation, the tongue’s center of gravity for eastern Japanese did not become more forward than western Japanese. As shown by Figure 4, eastern Japanese average rest position tended to be low and back. In contrast, western Japanese average rest position tended to be high and forward.

5 Conclusions and future work

In conclusion, we could find the difference of articulatory setting for eastern and western Japanese. First, we found western Japanese tongue’s center of gravity was higher and more forward than eastern. Furthermore, we found a relationship between the tongue’s rest position and each preceding vowel position. For both eastern and western Japanese, rest position is susceptible to the forward and backness of the preceding vowel position. However, it is insusceptible to the vowel height.

In future, I would like to do research comparing Japanese articulatory settings in a more limited area based on the results of this research, and look at the articulatory setting of lips, which was beyond the scope of this research.

Acknowledgements

First, I would like to express deep gratitude to Professor Ian Wilson, my supervisor, for all his help. Professor Yuichi Yaguchi advised about rotating the images. Naoya Horiguchi advised about my research and made the dot tracking program. I also thank all students who participated in data collection. Thanks to their help, my research was achieved.

References