Metrical Structure and Jaw Opening

Donna Erickson
Department of International Cultural Studies
Gifu City Women’s College, Gifu, Japan
erickson@gifu-cwc.ac.jp

Abstract

Acoustic and articulatory measurements were made of utterances containing seven syllables using the vowel /a I/ with varying degrees of prominence, including contrastive emphasis/focus. A trained phonetician judged the amount of prominence for each of the seven syllables. The findings show a significant correlation between amount of jaw opening for each of the target syllables and perceived prominence. There is also similarity between the patterns of perceived prominence for each sentence and the pattern of jaw opening for each utterance; however, more striking is the finding of alternating amounts of jaw opening for each syllable, suggesting that jaw opening may play an important articulatory role in implementing the foot-structure of an utterance.

1. Introduction

Previous studies on American English prosody have reported that the jaw tends to open more for syllables with greater prominence (e.g., [1, 2]). The height of the vowel also affects the amount of jaw opening involved in producing each syllable: high vowels require less open jaws and low vowels, more (e.g., [3]). However, [4] suggests a simple approach for normalizing vowel height in order to treat the prominence component of jaw opening separately from that of the vowel height component. It was proposed that articulation of American English prosody is closely connected to the pattern of jaw movement; specifically, the jaw sets the framework for the prominence characteristics of the syllable, and the tongue moves appropriately in order to produce the desired vowel.

In order to examine more carefully the hypothesis that the amount of syllable prominence of an utterance is reflected in the amount of jaw opening for the syllable, this paper examines the jaw movement production by an American English speaker producing utterances containing seven syllables using the vowel /a I/ with varying degrees of prominence, including contrastive emphasis/focus.

2. Methods

Articulatory and acoustic data were recorded at the University of Wisconsin X-Ray Microbeam Facilities, Madison, Wisconsin (e.g., [5]) from one American English male speaker (Midwest dialect, Wisconsin.) The sentences were “No, I saw five bright highlights in the sky tonight” in which the contrastive nuclear accent was assigned to either “five,” “bright,” or “highlights,” and “Yes, I saw five bright highlights in the sky tonight” with no contrastive nuclear accent. Each utterance was repeated 10-12 times, for a total of 36 utterances.

For recording articulatory movements, spherical gold pellets (2.4-3 mm in diameter) were glued onto selected points on the tongue, lips, and jaw of the speaker. Two pellets were attached to the mandible, one at the lower incisors, and one on the buccal surface of a lower molar tooth. For this data analysis, only the y-movement of the pellet attached to the mandible incisor (45 samples/s) was measured. In addition, reference pellets were affixed midsagittally to the nose bridge and to the center anterior surface of the maxillary incisors. These were used to determine the head coordinate system and set the maxillary occlusal plane at each instance during utterances (see [6], for more detailed descriptions). The x-axis corresponds to the intersection of the midsagittal plane and the maxillary occlusal plane, with the origin of the y-axis corresponding to the lowermost edge of the maxillary incisor. The y-axis is normal to the maxillary occlusal plane, intersecting the plane at the origin. Jaw opening is measured in terms of the vertical value of the jaw displacement from the maxillary occlusal plane. It is referred to as maximum jaw displacement, or simply jaw opening for the syllable.

Measurements of jaw x-y positions and LPC-Cepstrum method formant extraction and F0 estimation using a MATLAB-based program (written by J. Dang) were made at the time of maximum jaw opening during the digit. Figure 1 shows a sample utterance, and the vertical lines indicate the point where measurements were made.

Fig. 1. Sample display of “No, I saw five bright HIGHLIGHTS in the sky tonight.” The top two windows show the horizontal and vertical tongue dorsum positions; the next windows, the horizontal and vertical mandible positions (MNIx and MNIy, respectively) in mm, and the bottom two windows the acoustic signal and spectrogram. The arrows and vertical lines show the point where the articulatory and acoustic measurements were made at the lowest jaw opening in the target syllable.
Perceptual judgments of the prominence of the target syllables were done by a trained phonetician (the author) based on a method described in [7]. In brief, this is a practical approach for teaching rhythm to non-native speakers of English. It is based on the hypothesis that American English rhythm results from different weightings assigned to syllables, based on “levels” of sentence structure. For analyzing the utterances in this study, the following levels were assigned to each of the target syllables. Each syllable was assigned a level of “1.” However, within a foot, the most prominent syllable was assigned a level of “2,” within an intermediate phrase, the most prominent syllable, a “3,” within the utterance, a “4,” and for contrastively emphasized syllables, a “5.” The perceptual judgments of prominence were made prior to measuring the articulatory data.

3. Results

The amount of jaw opening for the seven target syllables was measured. An ANOVA with word emphasis (whether the word was contrastively emphasized) as the independent variable and vertical jaw position as the dependent variable was done. The results showed a significant difference in jaw position as a function of contrastive emphasis. In addition, a Pearson correlation analysis was done. The results show a significant correlation (p<0.01) between vertical jaw position and syllable prominence ratings.

Figures 2-5 show the averaged amount of jaw opening for each of the seven target syllables (“I,” “five,” “bright,” “high,” “light,” “sky,” “night”) in each of the utterances with the word “five,” “bright,” “highlights,” contrastively emphasized and no contrastive emphasis, respectively. The y-axis represents the amount of jaw opening (mm), and the x-axis, the target syllables. The larger the bar height, the greater the jaw opening for that syllable.

Figure 2. Amount of jaw opening (mm) for each of the target syllables in the sentence “No, I saw FIVE bright highlights in the sky tonight.” “FIVE” is represented by the 2nd bar in the graph.

Figure 3. Amount of jaw opening (mm) for each of the target syllables in the sentence “No, I saw five BRIGHT (3rd bar) highlights in the sky tonight.”

Figure 4. Amount of jaw opening (mm) for each of the target syllables in the sentence “No, I saw five bright HIGH (4th bar)LIGHTS in the sky tonight.”

Figure 5. Amount of jaw opening (mm) for each of the target syllables in the sentence “Yes, I saw five bright highlights in the sky tonight.”

Figure 3. Amount of jaw opening (mm) for each of the target syllables in the sentence “No, I saw five BRIGHT (3rd bar) highlights in the sky tonight.”
The above figures show that the averaged amount of jaw opening varies for each of the seven target syllables even though the vowel height in each of the syllables does not. What seems to vary is the position in the sentence and the prominence of the syllables.

A hypothesis to be investigated is to what extent the amount of jaw opening for each syllable reflects the metrical structure of the utterance. Figures 6-9 below show the amount of perceived prominence as rated by a trained phonetician, using the practical method described in [7]. The bars indicate the prominence values of each target word with an assigned value ranging from 1-5, “5” indicating contrastively emphasized prominence.

![Figure 6](image6.png)  
Figure 6. Perceived syllable prominence for each of the target syllables in the sentence “No, I saw FIVE (2nd bar) bright highlights in the sky tonight.” (Compare with Fig. 2)

![Figure 7](image7.png)  
Figure 7. Perceived syllable prominence for each of the target syllables in the sentence “No, I saw five BRIGHT (3rd bar) highlights in the sky tonight.” (Compare with Fig. 3)

![Figure 8](image8.png)  
Figure 8. Perceived syllable prominence for each of the target syllables in the sentence “No, I saw five bright HIGH(4th bar) LIGHTS in the sky tonight.” (Compare with Fig. 4)

![Figure 9](image9.png)  
Figure 9. Perceived syllable prominence for each of the target syllables in the sentence “Yes, I saw five bright highlights in the sky tonight.” (Compare with Fig. 5)

Visual comparisons of the actual measured jaw openings (Figures 2-5) with the perceived prominence ratings (Figures 6-9) indicate that there are similarities between the jaw opening values and the prominence ratings. Especially this seems to be so for those utterances in which “five” (figure 2) was contrastively emphasized (compare figure 2 with figure 6). However, it is not a perfect match. For instance, in figure 7, the perceived syllable prominence is largest for the emphasized syllable “bright,” but in figure 3, the jaw opening on the word “bright,” which is the nuclear contrastively emphasized word, is not the largest amount of jaw opening in the utterance; actually, the jaw opening on the word “five” is larger than on “bright.” Nevertheless, the amount of jaw opening on “bright” in figure 3 (in which “five” is contrastively emphasized) is larger than that in the other three figures in which “five” is NOT contrastively emphasized.
An ANOVA with perceived syllable prominence (ratings of 1-5) as the independent variable and vertical jaw position as the dependent variable was done. The averaged jaw opening values for each assigned prominence level are shown in Figure 10 below. A Bonferroni adjustment pairwise comparison of probabilities indicated that jaw opening values of prominence level 1 are significantly different from all other levels ($p<0.01$); those of level 2 are significantly different ($p<0.01$) from all other levels except level 4; those of level 3 are significantly different ($p<0.01$) from all levels except level 4, and those of level 5 except for level 4 are significantly different from all other levels ($p<0.01$) but between level 3 and level 5, the significance is $p<0.05$.

![Figure 10](image.png)

This suggests that prominence levels 4 and 5 (representing utterance-level prominence and contrastive emphasis, respectively) might not be that different in terms of jaw opening patterns; nor are the levels 3 and 4 (phrasal-level prominence and utterance-level prominence) that different in terms of jaw opening.

What is striking, however, with regard to the patterns of jaw opening in these utterances is the alternating levels of jaw opening: (low) (high-low) (high-low) (high-low).

### 4. Discussion

The results show that jaw opening varies according to position and prominence of the syllable within an utterance. Specifically, contrastively emphasized syllables have larger jaw opening than the other non-contrastively emphasized syllables. Moreover, there is a significant correlation between perceived prominence and amount of jaw opening.

There seems to be a relationship between metrical structure and jaw opening, but it is difficult to characterize it precisely. The striking observation from these results is the alternating pattern of s-w jaw movement, which may reflect the underlying foot structure of the utterance. When a syllable in a weak position receives increased prominence, the amount of jaw opening increases.

This study is an attempt to understand the articulatory organization of prosody. It appears that the jaw is an important articulator of prominence in American English, but more research is needed in order to better understand the articulatory underpinnings of prominence and metrical structure.

### 5. References


