MAXIMUM JAW DISPLACEMENT IN CONTRASTIVE EMPHASIS

Donna Erickson[1] and Osamu Fujimura[2]

[1]Center for Cognitive Science, The Ohio State University
[2]Department of Speech and Hearing Science, The Ohio State University
Columbus, Ohio 43221

ABSTRACT

Maximum jaw displacement for a syllable (or word), measured as the lowest vertical jaw position relative to the maxillary occlusal plane using x-ray microbeam data, was examined in the context of sets of utterances differing in terms of contrastive emphasis. It was found that there is a consistent and significant increase in jaw opening on the emphasized word for the three speakers examined. Moreover, the jaw opening on the other words in the utterance is also affected by emphasis according to this preliminary study: The amount of jaw opening of the syllables following emphasis is reduced, and the amount of drop in jaw opening from the emphasized syllable to the following syllable is increased.

1. INTRODUCTION

Numerous studies have reported changes in jaw movement patterns, e.g., jaw height, displacement, duration and velocity, associated with prosodic changes of an utterance [see e.g., 1, 2, 3, 4, 5, 6]. Focusing specifically on contrastive emphasis, [7, 8], contrastively emphasized words are reported to have more jaw opening than unemphasized words, in terms of the amount of vertical jaw displacement as measured from the maxillary occlusal plane. The finding of increased jaw displacement is consistent with acoustic studies showing increased duration and intensity for prominence [e.g., 9]. A recent acoustic study [10] suggests that emphasis is also a phrase level phenomenon involving a temporal rearrangement of all the words in an utterance, not just the word receiving emphasis.

We hypothesize that emphasis may affect jaw movement both in terms of increased opening on the word emphasized and also on the amount of opening on the other words in the utterance. The questions addressed in this paper are (1) does the measure *maximum jaw displacement*, defined as the lowest vertical jaw position relative to the maxillary occlusal plane, consistently and significantly increase with contrastive emphasis, and (2) is there an effect on the jaw displacement of other words in the utterance containing contrastive emphasis.

2. METHODS

Acoustic and articulatory recordings were made using the x-ray microbeam facilities at the University of Wisconsin [11], analyzing data initially collected by Wesbury and Fujimura [7] . Three American English subjects (1 man, 2 women) produced the question-answer sentences like "Is it 599 Pine Street? No, it's 59FIVE Pine Street," reading from a monitor display with a marking on the digit to be emphasized either in initial, middle, final position of the answer part following "no" or with no emphasis The three speakers generally spoke these sentences 80 times: 40 times on the 5 9 5-type utterances (10 times with no contrastive emphasis and 10 times each with contrastive emphasis on the initial, middle and final digit); and 40 times on the 9 5 9-type utterances (10 times with no contrastive emphasis and 10 times each with contrastive emphasis on the initial, middle and final digit).

Informal perception tests indicated that the placement of contrastive emphasis was perceived clearly as that intended with a corrective emphasis on the designated word, and that the "yes statements" (with no-emphasis) were also perceived as not containing any correction when the "yes" part was removed. These utterances spoken with no contrastive emphasis will be referred to as "reference utterances."

Utterance type 1:
Is it 5 9 5 Pine Street? Yes, it's 5 9 5 Pine Street.
Is it 9 9 5 Pine Street? No, it's FIVE 9 5 Pine Street.
Is it 5 5 5 Pine Street? No, it's 5 NINE 5 Pine Street.
Is it 5 9 9 Pine Street? No, it's 5 9 FIVE Pine Street.
Utterance type 2:
Is it 9 5 9 Pine Street? Yes, it's 9 5 9 Pine Street.
Is it 5 5 9 Pine Street? No, it's NINE 5 9 Pine Street.
Is it 9 9 9 Pine Street? No, it's 9 FIVE 9 Pine Street.
Is it 9 5 5 Pine Street? No, it's 9 5 NINE Pine Street.

Table 1: Corpus

To analyze the pellet traces (time functions) of the jaw movement data, we first visually examined the movement and speech signals, using the University of Madison Microbeam Facilities Display and x_y tool programs; then we measured the maximum vertical displacement, using a Macintosh-based software tool [12, 13]. This program, μTracker, automatically segments and measures movement segments of kinematic time functions in one selected articulatory dimension independently from acoustic input and from other articulatory dimensions.

Maximum vertical displacement was measured from the maxillary occlusal plane to the maximum opening during the vocalic gesture.

3. RESULTS

Measurements of the lowest vertical jaw position during each of the syllables are shown in Figures 1 and 2, for "5 9 5 Pine Street" and "9 5 9 Pine Street" utterances respectively .

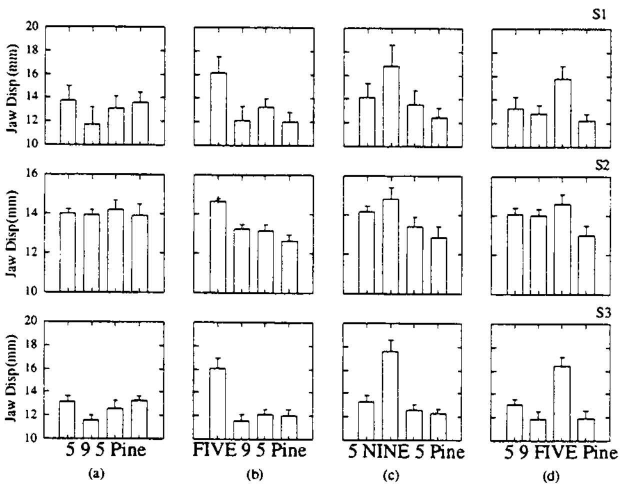

Figure 1: Measurements of the maximum vertical jaw displacement from the maxillary occlusal plane for "5 9 5 Pine (Street)" utterances for three speakers, shown in upper, middle and bottom rows. The y-axis represents the mean of the maximum vertical jaw displacement over 9 or 10 tokens; the word sequence in each utterance is specified in the x-axis. (a) reference utterance, (b) initial digit emphasized, (c) middle digit emphasized, and (d) final digit emphasized

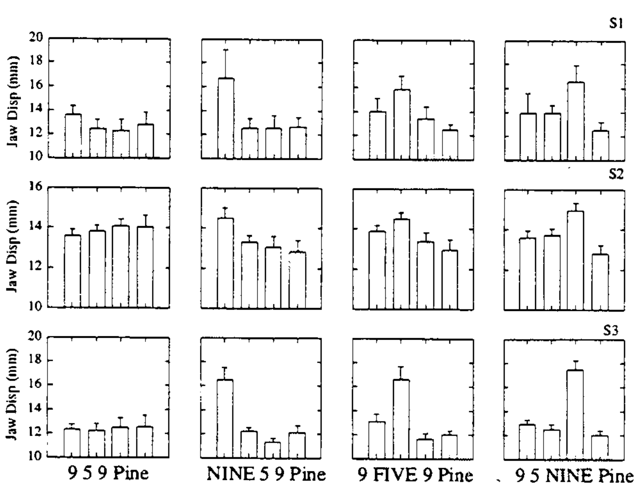

Figure 2: Measurements of the maximum vertical jaw displacement from the maxillary occlusal plane for "9 5 9 Pine (Street)" utterances for three speakers, shown in upper, middle and bottom rows. (Same format as Figure 1.)

For each of all three speakers, for both utterance types, the digit that receives the emphasis has the largest amount of jaw opening ($p < .001$) --even for speaker 2, who shows relatively little jaw movement compared to the other speakers. This confirms previous results by Westbury and Fujimura [7].

The pattern in Figures 1 and 2 reflects a general prosodic pattern of this type of utterance, as well as the effects of contrastive emphasis. In order to see directly the differential effects of emphasis, we subtracted the amount of jaw displacement in the reference utterances from that of the corresponding word in the utterances with emphasis. We used a Macintosh shareware program, "Binner," and produced "averaged" reference utterances for each digit sequence type, with mean jaw displacement values for each of the three digits plus "pine" in the reference utterances. These averaged jaw movements were subtracted from the individual jaw displacement values for each word in each of the emphasis conditions.

For all conditions for each speaker, the vertical jaw displacement is significantly greater for words intended with contrastive emphasis than for the corresponding words not intended with emphasis in the reference utterance ($p < .0001$, using a General Linear Model Procedure). The means are shown in column A of Table 2 below, along with the means (and standard deviations) for each speaker pooled over position and utterance type. There is a complex interaction between utterance type, position, and emphasis, which we are currently investigating. As can be seen in the figures above, for instance, the word in middle position generally shows greater jaw displacement than that in initial or final position.

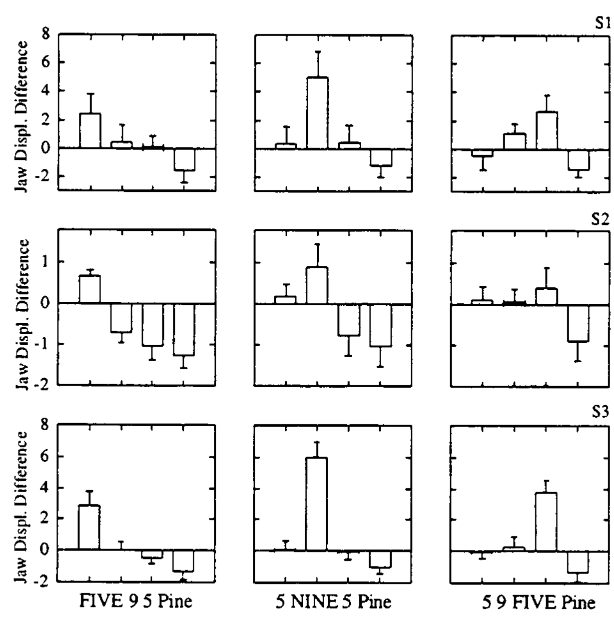

Figure 3: Bar graph displays of the *difference* in jaw displacement between the emphasized and unemphasized versions, same data as in Figure 1. The zero-line represents the amount of jaw opening for the words in the reference

utterance. The values above or below the zero line indicate the extent to which the jaw opening for the digits in the utterance with emphasis are greater or lesser, respectively, than the digits in the reference utterance.

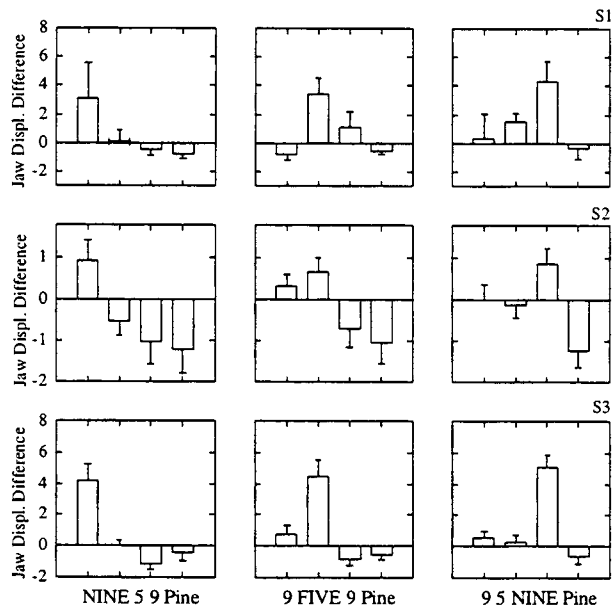

Figure 4: Bar graph displays of the *difference* in jaw displacement between the emphasized and nonemphasized versions, same data as in Figure 2. (Same format as Figure 3.)

There is a tendency for the syllable immediately following the emphasized digit to have a reduction in jaw opening, relative to the reference utterance, regardless of position or utterance type. This is especially clear for Speaker 2 (the non-jaw mover). Pooling the data across emphasized digit position, speakers and utterance types shows a significant reduction in the amount of jaw opening on the digit following the emphasized digit relative to the base-line ($p < .0079$). Moreover, there is a significant increase in amount of drop from emphasized digit to following unemphasized digit, ignoring the jaw closing action for the intervening consonants ($p < .0001$). The means are shown in columns B and C of Table 2 below, along with the means (and standard deviations) for each speaker pooled over position and utterance type.

	A. Emphasized word	**B. Following word**	**C. Drop from emphasized**
All	2.63 (p<.0001)	-.36 (p<.0079)	3.23 (p<.0001)
S1	2.84 (SD 2.07)	.18 (SD 1.19)	3.44 (SD 1.79)
S2	.74 (SD .46)	-.81 (SD .45)	1.54 (SD .61)
S3	4.41 (SD (1.32)	-.47 (SD .69)	4.88 (SD 1.28)

Table 2. Means of maximum jaw displacement of words in emphasized utterances relative to jaw displacement of words in reference utterance, pooled over all conditions (GLM) (top row) and by Speaker (T-test) (bottom three rows). Column A indicates the amount of extra jaw displacement observed for the emphasized word; column B indicates the amount of jaw displacement on the word following emphasis; and column C indicates the amount of drop in jaw displacement from the emphasized word to the following word.

4. DISCUSSION

In addition to the very local cue of emphasis as an increase of jaw displacement on the emphasized word, there is also a global cue: reduction of jaw opening for the immediately following word and a significant drop in jaw displacement from the emphasized to the following non-emphasized digit. Although we have not done a formal regression analysis, visual inspection of the data in Figures 3 and 4 suggests that the presence of emphasis affects the slope of the regression line of the jaw opening among the digits following the emphasized digit.

A comment about individual differences across speakers: Among the three speakers we have analyzed, Speakers 1 and 3 tend to show more similar patterns than does Speaker 2. For Speaker 2 the mean increase in maximum jaw movement of the emphasized word, compared to the same word in the reference utterance, (averaged over all conditions), is about 1 mm, as compared to that of Speakers 1 and 3 who show an averaged mean increase in maximum jaw movement of about 4 mm. However, Speaker 2 shows considerable reduction of jaw opening on the digit following the emphasized digit, as clearly seen in Figures 3 and 4 (note the ordinate scale is expanded for Speaker 2). Thus, although Speaker 2 does not show large increase on the emphasized digit as the other speakers do, she shows comparable reduction on the digit following emphasis, so that looking at the drop in jaw opening from the initial emphasized word to the following unemphasized word in Figures 3 and 4, the total amount of drop in jaw opening following emphasis is about the same (2 mm) as that seen for the other two speakers . If we look at the mean values of drop in jaw displacement averaged over position and utterance type (Table 2, column C), we see that for Speaker 2 the mean drop is 1.5 mm, about twice the size of the amount of "extra" jaw displacement given to the emphasized word. For Speakers 1 and 3, the amount of "extra" jaw displacement for the emphasized word by itself seems to account for most of the "drop." This suggests that speakers may be trading off cues. Speaker 2 seems to use the global cue of abrupt reduction in jaw movement for the contiguous (stressed) syllable nuclei more predominantly than the local cue of increased jaw movement on the emphasized word to signal emphasis.

In summary, emphasis affects jaw movement on both a local and global level. There is a direct effect of enhancement of jaw opening of the emphasized word, and an indirect effect on the word following the emphasis: the amount of jaw opening of the word following emphasis is reduced and concomitantly, the amount of drop in jaw opening from the emphasized word to the following word is increased. Individual differences across speakers suggests that speakers may be trading cues, with one speaker using the global cue more predominantly than the local cue for contrastive emphasis.

Future studies are planned to examine the jaw movement patterns of additional speakers on high vowels as well as low vowels. Also, we plan to investigate the regression lines of the jaw opening before and after emphasis in order to further explore the global effect of emphasis on an utterance. The interaction between position and utterance type, as well as interspeaker differences, will also be investigated in more detail.

Both the local and global effects of contrastive emphasis on jaw displacement may be accounted for by assuming an augmentation of the prosodic structure, e.g., a metrical grid, and also alteration of the metrical structure reflecting an altered phrasing of the same linguistic form. The C/D model proposes a scheme for such prosodic modification of an utterance. This issue is discussed elsewhere [14].

ACKNOWLEDGMENTS

Work supported by NSF SBR-951199B and by a gift to Osamu Fujimura from ATR/ITL, Kyoto, Japan. We thank Kevin Lenzo for his invaluable help in early stages of this paper.

REFERENCES

1. Summers, W. V. (1987). Effects of stress and final consonant voicing on vowel production: Articulatory and acoustic analyses. *J. Acoust. Soc. Am.*, **82**, 847-863.

2. Beckman , M. E., Edwards, J., and Fletcher, J. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In G. Docherty and D. R. Ladd (eds.) *Papers in Laboratory Phonology II: Segment, Gesture,, Prosody* (pp. 68-86). Cambridge: Cambridge University Press.

3. Edwards, J., Beckman, M., and Fletcher, J. (1991). The articulatory kinematics of final lengthening. *J. Acoust. Soc. Am.*, **89**, 369-382.

4. de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *J. Acoust. Soc. Am.*, **97**, 491-504.

5. de Jong, K. Beckman, M. E., and Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech*, **36**, 197-212.

6. Harrington, J., Palethorpe, S., Fletcher, J. & Beckman, M.E. (1996). Competing hypotheses concerning the articulation of stress in English. *J.Acoust.Soc.Am.*, **99** (A), 2494.

7. Westbury, J., and Fujimura, O. (1989). An articulatory characterization of contrastive emphasis. *J.Acoust.Soc.Am.*, **85** (Suppl.1), S98.

8. Erickson, D. and Honda, K. (1996). Jaw displacement and F0 in contrastive emphasis. *J.Acoust.Soc.Am.*, **99** (A), 2494.

9. Lehiste, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT Press.

10. Erickson, D. and Lehiste, I. (1995). *Proc. XIIIth Int Cong.Ph.Sc.*, **4**, 352-355.

11. Nadler, R., Abbs, J. H. and Fujimura, O. (1987). Speech movement research using the new x-ray microbeam system. In *Proc. XIth International Congress of Phonetic Sciences*, **1**, (pp. 221-224). Tallin, Estonia, USSR

12. Erickson, D., Lenzo, K. and Fujimura, O. (1994). Manifestations of contrastive emphasis in jaw movement. *J. Acoust. Soc. Am.*, **95**, (A), 2822 .

13. Lenzo, K., and Erickson, D. (1994). Measuring articulatory tract motion. *J. Acoust. Soc. Am.*, **95**, (A), 2818.

14. Fujimura, O. and Erickson, D. (1996). Prosodic organization of speech signals: Recent development in the C/D Model. *Presented at Sigphon, Santa Cruz, CA., June 1996.*