



Published in final edited form as:

J Speech Lang Hear Res. 2006 December ; 49(6): 1368–1379. doi:10.1044/1092-4388(2006/098).

Effects of Increasing Sound Pressure Level on Lip and Jaw Movement Parameters and Consistency in Young Adults

Jessica E. Huber, Ph.D. and Bharath Chandrasekaran, M.S.

Purdue University, Department of Speech, Language, and Hearing Sciences, West Lafayette, IN

Abstract

Purpose—Examination of movement parameters and consistency has been used to infer underlying neural control of movement. However, there has been no systematic investigation of whether the way individuals are asked (or cued) to increase loudness alters articulation. The aim of the current study was to examine whether different cues to elicit louder speech induce different lip and jaw movement parameters or consistency.

Methods—Thirty healthy young adults produced two sentences 1) at comfortable loudness, 2) while targeting 10 dB SPL above comfortable loudness on a sound level meter, 3) at twice their perceived comfortable loudness, and 4) while multi-talker noise was played in the background. Lip and jaw kinematics and acoustic measurements were made.

Results—Each of the loud conditions resulted in a similar amount of SPL increase, about 10 dB. Speech rate was slower in the background noise condition. Changes to movement parameters and consistency (relative to comfortable) were different in the targeting condition as compared to the other loud conditions.

Conclusions—The cues elicited different task demands, and therefore, different movement patterns were utilized by the speakers to achieve the target of increased loudness. Based on these results, cueing should be considered when eliciting increased vocal loudness in both clinical and research situations.

INTRODUCTION

Examination of speech movement parameters and consistency has been used extensively in the study of motor control to infer underlying neural control processes. In articulatory kinematic analyses, measures of movement parameters (duration, displacement, and velocity) and variability (standard deviations, covariance, and the spatiotemporal index (STI)), have been utilized to study normal speech production, changes to speech production with development or normal aging, and changes to speech production in individuals with motor disorders (Ackermann, Hertrich, Daum, Scharf, & Spieker, 1997; Ackermann, Hertrich, & Scharf, 1995; Dromey, Ramig, & Johnson, 1995; Forrest, Weismer, & Turner, 1989; Green, Moore, Higashikawa, & Steeve, 2000; Kleinow, Smith, & Ramig, 2001; McClean & Tasko, 2003; Perkell, Matthies, Svirsky, & Jordan, 1993; Schulman, 1989; Smith, Goffman, Zelaznik, Ying, & McGillen, 1995; Tasko & McClean, 2004; Wohlert & Smith, 1998).

Previous studies have shown that increasing loudness affects articulatory movement parameters. Lip and jaw opening displacement have been shown to increase with increases

in loudness (Schulman, 1989; Tasko & McClean, 2004). Increasing loudness has also been shown to result in increased first formant frequency for [α], most likely due to increased jaw opening and a resulting lower tongue position (Huber, Stathopoulos, Curione, Ash, & Johnson, 1999). The variability of the movement trajectories in loud speech has not been shown to change in young, healthy adults. In one study, movement consistency for an entire phrase, as measured by the STI, did not change significantly as the young adults spoke louder (Kleinow et al., 2001). However, there has been no systematic investigation of whether the manner in which individuals are cued to increase loudness alters articulatory movement parameters or consistency.

In previous studies which included loudness manipulations, several different cues have been used including asking participants to target a specific sound pressure level (SPL) using an SPL meter (Stathopoulos & Sapienza, 1997), asking participants to speak at twice or four-times their comfortable loudness (Dromey & Ramig, 1998; Kleinow et al., 2001), and presenting noise through headphones to participants as they spoke (Winkworth & Davis, 1997). In the current study, we chose to use cues similar to the ones used in previous studies employing loudness manipulations. Studying cues such as targeting a specific SPL using an SPL meter and asking participants to speak at twice their comfortable loudness is useful from a therapy standpoint since these types of cues are used in speech therapy aimed at increasing vocal loudness (c.f., Ramig, Countryman, Thompson, & Horii, 1995). Even though noise is not used in therapy, we chose to it as a cue since it is a more natural cue to speakers to increase vocal loudness. Speaking in noise has been shown to elicit an automatic increase in vocal loudness, without an overt cue (Lane & Tranel, 1971; Pick, Siegel, Fox, Garber, & Kearney, 1989). Further, speaking in noise is a familiar task for most individuals.

Many potential differences exist between instructing a participant to speak louder and asking a participant to speak in a noisy environment which will automatically trigger an increase in loudness. Instructing participants to speak louder will cause them to consciously focus on their loudness level, more so than they would in a natural situation such as in noise when louder speech is required in order to avoid communicative breakdown.

The task and the difficulty of the task may be perceived differently depending on the cue. For example, speaking at 85–90 dB SPL, as participants typically were in the SPL targeting condition in the current study, may have been perceived as harder than speaking twice as loud as their comfortable loudness level. This may be due to the fact that the participant's attention is drawn to the magnitude of loudness increase in the targeting condition. The participants may not realize how much louder they are speaking in the twice as loud and noise conditions.

The three cues chosen for the current study also differ in how feedback regarding loudness is provided. The way feedback is presented may alter the difficulty of the task or change the demands of the task. In the targeting condition, participants were given visual feedback from the SPL meter. In the twice as loud and noise conditions, participants used their own auditory feedback to judge if they were speaking loud enough. In the noise condition, the auditory judgment was more difficult due to the presence of the noise and may also have involved more than just a loudness target. For example, individuals also speak more slowly, possibly to improve intelligibility, in noise (c.f., Huber, Chandrasekaran, & Wolstencroft, 2005; Van Summers, Pisoni, Bernacki, Pedlow, & Stokes, 1988). Conversely, Dromey and Ramig (1998) did not find a change in speaking rate when participants were asked to speak at twice or four-times as loud as comfortable.

Previous studies have demonstrated that changes to task demands can result in differences in the movement parameters and consistency of movement. Gentilucci, Benuzzi, Bertolani,

Daprati, and Gangitano (2000) demonstrated that word reading has a greater effect on movement than the initial percept of the object. When large objects were labeled with the word “small,” reaching and grasping movements were more like those targeting small objects than large objects. In this case, changes in the perceived task demands due to the labels on the objects resulted in differences in reaching and grasping movements. Tasko and McClean (2004) demonstrated that task demands result in differences in articulator movements. They found that speed, displacement, duration, and variability of articulator movements were different when participants said a nonsense phrase (“a bad daba”) as compared to a sentence, reading passage, and monologue.

Changes to the task demands as a result of different cues to increase loudness would be reflected in how the movement parameters are modified to achieve louder speech. Along these lines, Huber et al. (2005) demonstrated that respiratory mechanisms differed depending on the cue used to elicit increased loudness. A natural hypothesis, given previous data on changing task demands from the limbs, articulators, and respiratory system, would be that different cues to increase loudness will alter lip and jaw kinematic parameters and/or consistency.

The purpose of the current study was to examine how task demands affect speech kinematics. Specifically of interest was whether different cueing methods to elicit louder speech induce different articulatory movement parameters or consistency. Lip and jaw kinematic measurements were used to examine lip and jaw movement and consistency. Formant frequency measurements were used to infer information about tongue placement and vocal tract shape. Studies which employ loudness changes serve two purposes. In demonstrating how the speech system functions outside the range of comfortable speaking conditions, these studies extend our current understanding of the neural control of the speech system (Tasko & McClean, 2004). Loudness change acts a kind of natural perturbation to the speech system since there are requirements for loud speech which are not present for speech at comfortable loudness. Further, since many therapeutic strategies with adults with neuromotor control diseases incorporate changes to loudness, these studies provide information related to the use of loudness therapies to compensate for speech difficulties.

METHODS

Participants

Thirty normal young adults, 15 women and 15 men, participated in the study. The mean age of the women was 22 years, 4 months, and the mean age of the men was 22 years, 10 months. Participants reported no history of voice problems, neurological disease, head or neck surgery, formal speaking or singing training; no recent colds or infections; and that they had been non-smoking for the past five years. They had normal speech, language, and voice, as judged by a certified speech-language pathologist (the first author). Participants were Purdue University students and staff from different areas of the United States; however, none of the participants had an accent which would be different from the predominant speech patterns in the Midwest. Participants had normal hearing as indicated by a hearing screening at 30dB HL for octave frequencies between 250 and 8000 Hz, bilaterally, completed in a quiet room.

Procedures and Speech Tasks

Participants said two sentences: 1) SHORT: “Buy Bobby a puppy” and 2) EMBED: “You buy Bobby a puppy now if he wants one.” We chose these two utterances in order to more fully sample the segments of interest (“Buy Bobby a puppy” and “Bob”). Participants were instructed to say one sentence per breath and to speak clearly and audibly to the

experimenters. Each sentence was said fifteen times consecutively in the following four conditions. Four cues (conditions) were chosen which replicated cues used in previous studies of changes to speech production with increased loudness.

1. Participants were instructed as follows: “read the sentences at your comfortable loudness and pitch” (COMF condition).
2. Participants were instructed to read the sentences at what they perceived as twice their comfortable loudness (2XCOMF condition). The instruction to the participants was to “read the sentence at what you feel is twice your comfortable loudness.” No feedback regarding the participant’s loudness was provided during this condition. Participants were offered an opportunity to practice attaining a level they felt was twice as loud as comfortable. If this condition did not directly follow the COMF condition, participants were first instructed to read the sentence at their comfortable loudness and pitch until the examiner saw their SPL level return to a level similar to their original SPL level for the COMF condition. This ensured that the “twice as loud as comfortable” loudness level was relative to the individual’s normal loudness level.

Participants were instructed target a specific SPL, using an SPL meter providing numerical feedback (COMF+10 condition). They were instructed as follows: “The number goes up as you get louder. When you read the sentence this time, I want you to keep that number between XX and XX.” The SPL targets were inserted in for the “XX.” The SPL targets were set at 10 dB above the participant’s comfortable SPL (+/- 2 dB). A target of 10 dB above comfortable was chosen since, based on previous data demonstrating that participants perceive an increase in SPL of 10 dB as a doubling in loudness (Stevens, 1955), we expected participants to increase SPL by 10 dB in the 2XCOMF condition. The SPL meter was set to C-weighting and fast response during data collection. Participants were offered an opportunity to practice attaining the specified SPL level prior to the start of data collection for this condition.

3. Participants were instructed to read the sentences while noise was played in the background (NOISE condition). In this condition, the noise was turned on and the participants were instructed to “read the sentence.” No feedback regarding the participant’s loudness was provided. None of the participants requested practice trials in this condition. The noise consisted of multi-talker babbling noise (AUDiTEC of St. Louis), played at 70 dBA relative to the participant’s ears. The speakers were placed in front of the participant, 39 inches away.

The fifteen trials of the SHORT sentence were produced first, followed by fifteen trials of the EMBED sentence in each condition. The COMF condition was always completed first. The order of the three loud conditions (COMF+10, 2XCOMF, and NOISE) was counterbalanced across participants.

Equipment

The acoustic signal was transduced via a condenser microphone which was connected to an SPL meter (Quest model 1700). The microphone was placed 6 inches from the participant’s mouth, at a 45 degree angle. The microphone signal was recorded to digital audiotape (DAT) and later digitized into a PC-computer using Praat (Boersma & Weenink, 2003). The signal was digitized at 44.1 kHz and resampled at 18 kHz. The resampling process applied a low-pass filter at 9000 Hz for anti-aliasing.

The microphone signal was calibrated using a piston-phone set to output a 94 dB signal of 1000 Hz. The calibration signal was read on the SPL meter to ensure that the meter was

functioning correctly for each participant. The calibration signal was also collected directly to the DAT tape prior to data collection for each participant. The calibration signal was digitized, along with the microphone signal for each participant, to the computer. The difference between the known value of the calibration signal (94 dB) and the actual SPL measured in Praat was applied to all SPL measurements.

The lip and jaw kinematic signals were transduced using infrared light emitting diodes and a camera system (Optotrak 3020 system, Northern Digital Inc.). Markers were attached to the skin surface near the vermilion border of the upper and lower lips at midline using surface EMG adhesive collars. The jaw marker was attached to a lightweight splint which was attached to the chin at midline using medical adhesive tape. The splint ensured that the jaw marker could be transduced by the camera system as the jaw moved. The marker on the lower lip reflected movements of the lower lip and the jaw. The marker on the jaw reflected movements of the jaw, and the marker on the upper lip reflected movements of the upper lip. Head motion was tracked using five iresds, one attached to the forehead at midline and four attached to specially-modified transparent sports goggles (Walsh & Smith, 2002). Two of the four iresds were attached to the goggles at the level of the lateral angle of the eye on the left and the right sides. The other two were attached to the goggles at the level of the angle of the mouth on the left and right sides. Data from these five markers were used to determine the three-dimensional axes of the head for each participant, and the movements of the upper and lower lips and jaw were calculated relative to the head coordinate system, allowing for the correction for any artifact resulting from head motion. Data from the iresds was digitized by the Optotrak system at 250 Hz. Movement of the articulators in superior-inferior dimension was analyzed since that is the primary dimension of movement for bilabial stops. An audio signal, digitized at 2000 Hz, was collected in synchrony with the lip and jaw kinematic data and used to identify utterances associated with kinematic events.

Measurements

The first two trials of each utterance in each condition were discarded. The next ten consecutive trials which were produced without error were chosen for analysis. Trials were discarded if words were added or missed or if disfluencies or hesitations were present.

Acoustic Measurements

1. Sound Pressure Level (SPL) was measured as the average across each sentence, including all parts of the utterance. Measurement of SPL provided information regarding the magnitude of change in vocal loudness in each of the loud conditions.
2. First Formant Frequency (F1) was measured for the steady-state portion of [α] in “Bobby” using a Fast Fourier Transform (FFT). The area of F1 was identified as the first peak resonance in the FFT spectrum. The standard settings for FFT analysis in Praat were used, including a Gaussian window with a window length of 5 ms. The frequency of the highest harmonic in the area of the first formant was taken as the F1. Measurement of F1 was included to provide some information regarding tongue placement and vocal tract configuration.
3. Second Formant Frequency (F2) was measured for the steady-state portion of [α] in “Bobby” using an FFT. The area of F2 was identified as the second peak resonance in the FFT spectrum. The same settings were used for the FFT, as for the measurement of F1. The frequency of the highest harmonic in the area of the second formant was taken as the F2. Measurement of F2 was included to provide some information regarding tongue placement and vocal tract configuration.
4. Duration of the sentence (SDUR) was measured as the time to produce the entire sentence. The onset and offset of the sentence were determined by the onset and

offset of voicing, as shown in the microphone signal. Measurement of SDUR was included to provide information regarding speech rate. Syllables per second were not used since the number of syllables did not change across conditions or participants.

Lip and Jaw Kinematic Measurements—Articulatory kinematic measurements were made using algorithms written to run within Matlab. Before any measurements were made, the articulatory kinematic signals were low-pass filtered at 40 Hz. The jaw signal (JW) was subtracted from the combined lower lip and jaw signal to compute the lower lip signal (LL).¹ The “Buy Bobby a puppy” portion of each utterance was segmented from the entire kinematic signal for the three kinematic traces, LL, JW, and upper lip (UL) (see Figure 1). In order to segment the utterance, the displacement signals were differentiated to obtain the velocity signals. The utterance was segmented at the maximum opening velocity for the vowel [ai] in “buy” and the maximum opening velocity for the vowel [i] in “puppy” (Smith et al., 1995) (see Figure 1, lines A and B). The spatiotemporal index (STI) was computed across the segmented portions of the ten productions of each sentence in each condition for each articulator (Smith, Johnson, McGillem, & Goffman, 2000).

Measurements were made from the production of “Bobby” in each utterance for each articulator. The segment “Bob” was chosen for analysis since it is characterized by a labial open-close movement, making the measurement of the lip and jaw kinematics more reliable. Also, “Bob” was consistently stressed in the sentences produced. The “pup” from “puppy” was not chosen for analysis because “puppy” came at the end of the SHORT sentence, and therefore, was not always completely produced or stressed. We wanted to choose a segment which would be produced similarly across participants, sentences, and conditions. Acoustic measurements were collected from “Bob” so that they could be interpreted with the lip and jaw kinematic measurements.

1. Duration of “Bob” (BDUR) was measured from the peak closing displacement for the first “B” to the peak closing displacement for the second “b” in “Bobby” (see Figure 1, point C to point E).
2. Opening Displacement (ODISP) and Velocity (OVEL) were measured as the peak displacement or velocity, respectively, for [ɑ] in “Bobby” (see Figure 1, displacement: difference from point C to point D; velocity: point F).
3. Closing Displacement (CDISP) and Velocity (CVEL) were measured as the peak displacement or velocity, respectively, for the second “b” in “Bobby” (see Figure 1, displacement: difference from point D to point E; velocity: point G).

Statistics

The data from the two sentences were averaged. For SPL, F1, F2, and SDUR measurements, means were computed for each participant for each condition. The differences in the means for these measures were assessed in one-factor repeated measures analyses of variance (ANOVAs). For the articulatory kinematic measurements, means were computed for each participant for each condition and articulator. For these measurements, the differences between means were assessed in two-factor repeated measures ANOVAs. The within factors were condition and articulator (for the articulatory kinematic measurements). Tukey HSD

¹Since the movement of the jaw involves rotational and translational movements, simple subtraction to decouple the lower lip from the jaw will result in some error in the LL signal (Westbury, Lindstrom, & McClean, 2002). This error has been reported to be larger for low vowels (such as [ɑ] in “Bob”) due to the lower jaw position and could be larger in loud speech since it is likely to result in a lower jaw position as well (Westbury et al., 2002).

tests were completed for all factors and interactions which were significant in the ANOVA. The alpha level for the ANOVAs and the Tukey HSD tests was set at $p = 0.01$.

Inter-measurer reliability was completed on 2 male and 2 female participants, randomly chosen. Mean differences between the two sets of measurements were as follows, SPL: 0.10 decibels (dB), F1: 0.28 Hertz (Hz), F2: 9.71 Hz, SDUR: 0.02 seconds (s), BDUR: 0.001 s, ODISP: 0.006 millimeters (mm), CDISP: 0.002 mm, OVEL: 0.05 mm/s, CVEL: 0.05 mm/s, and STI: 0.11. Independent samples *t*-tests were computed between the first and second measurement for each variable. None of the alpha levels neared significance, ranging from $p = 0.194$ to $p = 0.992$, indicating good inter-measurer reliability.

RESULTS

Means and standard deviations for the conditions are presented in Table 1. Means and standard deviations for the articulators are presented in Table 2. The statistical results are presented in Table 3. Effect sizes (*d*) for the significant pairwise comparisons are given in the text.

Sound Pressure Level (SPL)

For SPL, there was a significant condition effect. The mean SPL was significantly higher for the three loud conditions than the COMF condition [COMF+10 $d = -2.38$; 2XCOMF $d = -1.95$; NOISE $d = -2.33$], but there were no significant differences in SPL among the three loud conditions.

First Formant Frequency (F1)

For F1, there was a significant condition effect. The mean F1 in the loud conditions was significantly higher than the mean F1 in the COMF condition [COMF+10 $d = -0.65$; 2XCOMF $d = -0.63$; NOISE $d = -0.79$], but there were no differences among the three loud conditions. As expected, the women ($M = 910.81$ Hz, $SD = 86.47$ Hz) had a higher F1 than the men ($M = 698.57$ Hz, $SD = 60.83$ Hz).

Second Formant Frequency (F2)

For F2, there was a significant condition effect. The mean F2 in the COMF+10 condition was significantly higher than the mean F2 in the COMF condition [$d = -0.45$], but there was no significant difference between the mean F2 for the COMF condition and the 2XCOMF and NOISE conditions. As expected, the women ($M = 1410.27$ Hz, $SD = 151.06$ Hz) had a higher F2 than the men ($M = 1169.00$ Hz, $SD = 96.00$ Hz).

Sentence Duration (SDUR)

For SDUR, there was a significant condition effect. The sentences produced in the NOISE condition were significantly longer in duration than those in the COMF condition [$d = -0.35$].

“Bob” Duration (BDUR)

For BDUR, there was a significant articulator effect, but no condition or interaction effects. BDUR was significantly longer in the JW than in the LL [$d = 0.68$] or the UL [$d = 0.95$]. BDUR was significantly longer in the LL than in the UL [$d = 0.34$].

Opening Displacement (ODISP)

For ODISP, there were significant condition and articulator effects and a significant articulator by condition interaction effect. The articulator by condition interaction is

depicted in Figure 2. In COMF, ODISP for LL was significantly larger than for JW [$d = -0.40$] and UL [$d = 2.41$], and ODISP for JW was significantly larger than for UL [$d = 1.36$]. For the three loud conditions, ODISP for LL and JW was significantly larger than for UL [COMF+10 LL $d = 2.50$, JW $d = 1.58$; 2XCOMF LL $d = 2.43$, JW $d = 1.79$; NOISE LL $d = 2.41$, JW $d = 1.67$], but there was no significant difference between the LL and JW. Opening displacement for both LL and JW increased significantly in all three loud conditions as compared to COMF [COMF+10 LL $d = -0.35$, JW $d = -0.54$; 2XCOMF LL $d = -0.50$, JW $d = -0.65$; NOISE LL $d = -0.52$, JW $d = -0.56$], but there were no significant differences among the three loud conditions for either articulator. There was no significant difference in ODISP for the UL across the four conditions.

Closing Displacement (CDISP)

For CDISP, there were significant condition and articulator effects and a significant condition by articulator interaction effect. The articulator by condition interaction is depicted in Figure 3. In COMF, CDISP for LL [$d = 2.30$] and JW [$d = 1.81$] were significantly larger than for the UL, but there was no significant difference in CDISP between LL and JW. In the three loud conditions, CDISP was significantly larger for the LL and JW than for UL [COMF+10 LL $d = 2.41$, JW $d = 2.02$; 2XCOMF LL $d = 2.00$, JW $d = 2.14$; NOISE LL $d = 2.30$, JW $d = 2.18$], and CDISP was significantly larger for JW as compared to LL [COMF+10 $d = 0.02$; 2XCOMF $d = 0.41$; NOISE $d = 0.33$]. The effect size for the comparison between LL and JW in COMF+10 is very small, suggesting this effect is not a real effect. Closing displacement was significantly larger for both the LL and JW in the three loud conditions as compared to COMF [COMF+10 LL $d = -0.35$, JW $d = -0.57$; 2XCOMF LL $d = -0.40$, JW $d = -0.66$; NOISE LL $d = -0.50$, JW $d = -0.63$], but there were no significant differences among the three loud conditions for either articulator. The CDISP for the UL did not change across the four conditions.

Opening Velocity (OVEL)

For OVEL, there were significant condition and articulator effects and a significant condition by articulator interaction effect. The articulator by condition interaction is depicted in Figure 4. In COMF, OVEL for LL [$d = -2.07$] and JW [$d = -1.30$] was significantly larger than for the UL, and OVEL for LL was significantly larger than for JW [$d = -0.25$]. In the three loud conditions, OVEL was significantly larger for LL and JW than for the UL [COMF+10 LL $d = -1.87$, JW $d = -1.35$; 2XCOMF LL $d = -2.00$, JW $d = -1.59$; NOISE LL $d = -2.04$, JW $d = -1.50$], but there was no significant difference between LL and JW. Mean OVEL for the LL was larger than in the 2XCOMF [$d = 0.41$] and NOISE [$d = 0.45$] conditions than in COMF, but there was no significant difference between COMF and COMF+10. For all three loud conditions, OVEL for the JW was larger than in COMF [COMF+10 $d = 0.46$; 2XCOMF $d = 0.60$; NOISE $d = 0.50$]. The OVEL for the UL did not change across the four conditions.

Closing Velocity (CVEL)

For CVEL, there were significant condition and articulator effects and a significant condition by articulator interaction effect. The articulator by condition interaction is depicted in Figure 5. In COMF, CVEL for LL [$d = 2.56$] and JW [$d = 1.63$] was significantly larger than for UL, and CVEL for LL was significantly larger than for JW [$d = 0.39$]. In the three loud conditions, CVEL was significantly larger for LL and JW than for the UL [COMF+10 LL $d = 2.73$, JW $d = 1.75$; 2XCOMF LL $d = 2.46$, JW $d = 1.82$; NOISE LL $d = 2.54$, JW $d = 1.85$]. In COMF+10 and 2XCOMF conditions, there was no significant difference in CVEL between LL and JW. In the NOISE condition, CVEL for the LL was significantly larger than for the JW [$d = 0.20$]. For all three loud conditions, CVEL was significantly larger for LL and JW than in COMF [COMF+10 LL $d = -0.38$, JW $d = -0.52$;

2XCOMF LL $d = -0.41$, JW $d = -0.59$; NOISE LL $d = -0.47$, JW $d = -0.51$]. The CVEL for the UL did not change across the four conditions.

Spatiotemporal Index (STI)

For the STI, there were significant condition and articulator effects, but no interaction effect. The STI was significantly lower for the COMF+10 condition as compared to the COMF condition [$d = 0.37$], but there was no significant difference among the 2XCOMF, NOISE, and COMF conditions. The STI was significantly higher for the UL [LL $d = -0.88$; JW $d = -1.33$], as compared to the other two articulators.

Summary of Loudness Related Changes

- SPL and F1 were higher in the three loud conditions than in COMF.
- F2 was higher in COMF+10 than in COMF.
- SDUR was higher in NOISE than in COMF.
- ODISP and CDISP for LL and JW were higher in the three loud conditions than in COMF.
- OVEL for LL was higher in 2XCOMF and NOISE than in COMF. OVEL for JW was higher in all three loud conditions than in COMF.
- CVEL for LL and JW was higher in all three loud conditions than in COMF.
- There was no change in UL displacement or velocity with changes in loudness.
- STI was lower in COMF+10 than COMF.

DISCUSSION

We examined whether different cues to speak louder resulted in different lip and jaw kinematics. For the most part, the type of cued used did not result in large changes to lip and jaw kinematics. However, differences in task demands in the NOISE cue resulted in altered speech rate and in the COMF+10 cue resulted in changes to some movement parameters and consistency. Further, the three articulators did not respond to the increase in loudness in the same way, indicating there are differences in the way in which the nervous system controls the various articulators to increase loudness.

Each of the conditions resulted in a similar level of SPL increase, about 10 dB. Therefore, any differences among the conditions are related to the type of cue used to elicit the increased loudness, rather than differences in the level of loudness increase among conditions.

The changes to sentence duration in the NOISE condition suggest that the NOISE cue elicited two task demands, increasing loudness and slowing speech rate. Sentence duration increased in the NOISE condition as compared to COMF, but not in the other two loud conditions. Participants may have spoken more slowly in the NOISE condition to improve transmission of the message in the background noise by increasing the salience of speech. The effect on loudness demonstrated in the current study has been demonstrated previously in studies of the Lombard Effect (c.f., Amazi & Garber, 1982; Lane & Tranel, 1971; Pick et al., 1989; Van Summers et al., 1988). Van Summers and colleagues also demonstrated that speakers reduce their rate of speech when speaking in noise and that speech produced in noise is more intelligible than speech produced in quiet, when the signals are presented at similar signal-to-noise ratios, possibly due to the reduction in speech rate.

There was no change in “Bob” duration, so it is likely that the change in sentence duration was due to an increase in pause time rather than articulatory time. An increase in pause time has been shown to improve intelligibility, most likely by allowing listeners more time to process the incoming speech signal, thereby increasing comprehensibility (Mendel, Walton, Hamill, & Pelton, 2003; Parkhurst & Levitt, 1978; Yorkston, Hammen, Beukelman, & Traynor, 1990). However, since sentences produced with overt hesitations were not included in the analysis, the increase in pause time could not have been large.

The changes in movement variability, second formant frequency, and opening velocity in the COMF+10 condition suggest that speakers constrained articulatory movement and reduced trial-to-trial movement variability. The COMF+10 condition resulted in a significant increase in movement consistency, as demonstrated by the STI, compared to the COMF condition; however, the other two loud conditions did not result in a change in STI. Normal adult speakers have highly patterned speech movements, but in the COMF+10 condition, the normal amount of variability present in articulatory movements was reduced. In this condition, the SPL meter provided external validation to the speaker as to when the target had been achieved. This was different from the other two loud conditions in which the target SPL was not specified and in which no external feedback was provided about a speaker’s loudness increase. The lack of normal variability suggests that the task demand of targeting a specific SPL level, as displayed by the meter, resulted in speakers programming the movement as consistently as possible from trial to trial. In other words, once they found the articulatory movements which allowed them to maintain the SPL in the appropriate range, they repeated those as exact movements as closely as possible in all productions of the sentences for the COMF+10 condition.

It is also possible that targeting a specific SPL on the SPL meter added a cognitive load. One way a speaker might respond to an increase in cognitive load is to reduce the degrees of freedom in their speech movement, thereby increasing the consistency of the movement. However, the effect of a cognitive load on lip and jaw movement consistency is unclear. Dromey and Bates (2005) found a slight, but nonsignificant, increase in STI in participants who were speaking while performing a mathematics task. Dromey and Benson (2003) found a significant increase in STI in participants who were speaking while performing a mathematics task, different from the one used by Dromey and Bates (2005). An increase in STI would suggest less consistency of movement with increased cognitive load.

The F2 data support the hypothesis that participants constrained their movements in the COMF+10 condition as compared to the other conditions. The second formant frequency can be affected by the amount of tongue constriction and the area of the front cavity (Fant, 1970). The opening displacement for the lower lip and jaw increased significantly for all three loud conditions. An increase in jaw opening would lead to less tongue constriction which would raise F2. In the COMF+10 condition, F2 did increase significantly, most likely due to the widening of the tongue constriction with increased jaw opening. However, in the 2XCOMF and NOISE conditions, the F2 did not change significantly. One plausible explanation for the lack of F2 change is that the tongue was also backed in the 2XCOMF and NOISE conditions, resulting in an exaggerated tongue posture for [α] (Huber et al., 1999). Backing the tongue would result in an increase in the size of the cavity in front of the constriction, which would lower F2. So the change in F2 resulting from tongue backing is opposite to the change which results from a larger area of constriction. Therefore no significant change in F2 would be present when an exaggerated tongue posture is used (backed and lowered tongue).

These data support that articulatory movements were constrained in the COMF+10 condition since the F2 does change, suggesting less exaggeration of the tongue posture for

[α] in the COMF+10 condition, as compared to the other two loud conditions. Additionally, for the COMF+10 condition, the mean increase in jaw opening displacement and F1 was smaller than the mean increases for the other two loud conditions, so participants, on average, did not increase jaw opening as much in response to the COMF+10 cue.

Changes in opening velocity demonstrate cue-related differences in lower lip movement which also suggest that movement was constrained in the COMF+10 condition. Opening velocity for the lower lip was larger in 2XCOMF and NOISE conditions than in COMF. However, opening velocity in the COMF+10 condition did not differ significantly from COMF, so the rate of lower lip movement was lower in the COMF+10 condition than in the other loud conditions. Rate of lower lip movement may have been less in this condition, as compared to the other loud conditions, to enhance control of the movement. It is not clear why this difference was not significant in opening displacement, although the trend for a smaller opening displacement in the COMF+10 condition is present.

In addition to demonstrating the differential effects of task demands on articulation, the current study demonstrated the displacement and velocity of movements for loud speech are different across the articulators, regardless of the cue. The movements of the upper lip did not change significantly with increased loudness. This suggests that the upper lip may not be as important to the mouth opening gesture associated with increased loudness as the lower lip and jaw. Since the upper jaw is stationary, the upper lip has less freedom to move in the superior-inferior dimension, and therefore, may not be in a position to participate in a large change in mouth opening.

Additionally, in all four conditions, movement of the upper lip was the most variable. Walsh and Smith (2002) demonstrated that the jaw was less variable than the lower and upper lips, but there was no difference between the upper and lower lips. A possible interpretation of the finding in the current study that the upper lip is more variable than both the lower lip and the jaw is that the movement targets for the upper lip for bilabial stops are less constrained than those for the jaw and lower lip. However, the variability of upper lip movement, as compared to the lower lip and jaw, may reflect the differences in bony architecture. The lower lip rides on the jaw, and the bony architecture of the jaw may constrain lower lip movement in the superior-inferior dimension, reducing variability.

Both the lower lip and the jaw contributed to increased mouth opening with increased loudness, as was demonstrated by the opening displacement data. However, in the 2XCOMF and NOISE conditions, the lower lip did not demonstrate as much closing displacement as the jaw. This finding may reflect saturation effects. The lower lip would be likely to stop moving once it contacts the upper lip, indicating full lip closure. The jaw may still move, compressing the lips together.

Some of the measurements reflected similarities in the neural control of the articulators. For example, the articulator by condition interaction for the STI measure was not significant in the current study, indicating that movement consistency for the articulators do not change differentially as a result of the type of cue used. Further, the finding that the 2XCOMF condition did not result in a change in movement variability, as measured by the STI, extends Kleinow and colleagues' (2001) findings. The current study was different from Kleinow and colleagues' study in that the current study examined the consistency of the lower lip, jaw, and upper lip separately, whereas Kleinow and colleagues' study used the composite lower lip plus jaw signal.

Finally, data from the current study demonstrated a relationship between F1 and jaw opening in loud speech. The first formant frequency increased all three loud conditions. This change in F1 was likely due to the increased jaw opening (Fant, 1970). Studies of lip and

jaw kinematics in loud speech have hypothesized that the increased jaw opening reported would result in an increase in F1 (Schulman, 1989). Also, studies of the acoustics of loud speech have hypothesized that increased F1 related to jaw opening (Huber et al., 1999). The present data, articulatory kinematics and acoustics from a single set of participants, verify the presence of this relationship. Since the F1 and jaw opening displacement condition effects were present in all three loud conditions, the relationship between jaw opening and F1 was present regardless of how the participant was cued.

In summary, movement parameters, overall trajectory variability, and resulting acoustics did change depending on how the speakers were cued to increase loudness. The slowing of speech rate in the NOISE condition suggested that the task demands elicited by the NOISE cue were to increase loudness and slow speech, most likely to improve intelligibility of speech in noise. The changes in the STI, F2, and opening displacement in the COMF+10 condition suggested that constrained movements were required to achieve the task demands in the targeting condition, leading to a reduction in the trial-to-trial variability usually seen in speech movements. The lower lip and jaw were the prime articulators for the mouth opening gesture associated with increased loudness and increased mouth opening was associated with increase F1, regardless of the cue used.

Acknowledgments

Grants

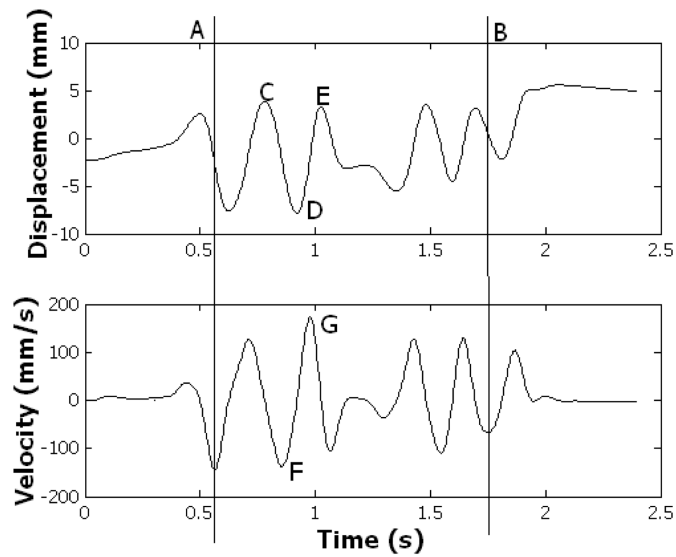
This research was funded by the National Institutes of Health, National Institute on Deafness and Other Communication Disorders, grant # 1R03DC05731.

REFERENCES

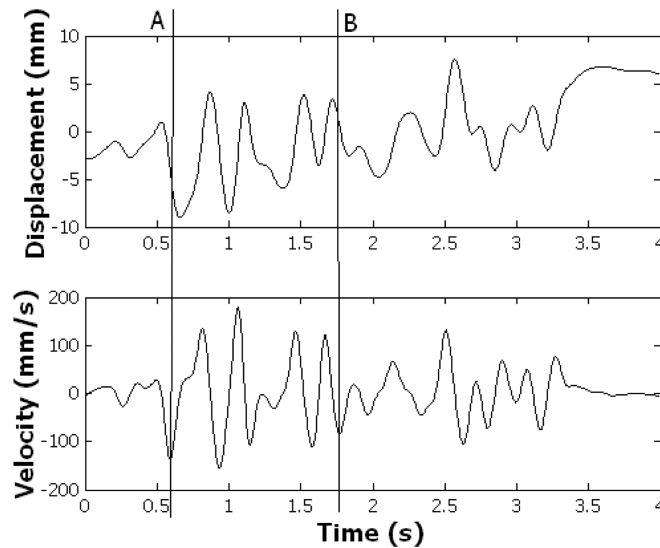
- Ackermann H, Hertrich I, Daum I, Scharf G, Spieker S. Kinematic analysis of articulatory movements in central motor disorders. *Movement Disorders*. 1997; 12(6):1019–1027. [PubMed: 9399230]
- Ackermann H, Hertrich I, Scharf G. Kinematic analysis of lower lip movements in ataxic dysarthria. *Journal of Speech and Hearing Research*. 1995; 38:1252–1259. [PubMed: 8747818]
- Amazi DK, Garber SR. The Lombard sign as a function of age and task. *Journal of Speech and Hearing Research*. 1982; 25:581–585. [PubMed: 7162159]
- Boersma P, Weenink D. Praat (Version 4.1). 2003
- Dromey C, Bates E. Speech interactions with linguistic, cognitive, and visuomotor tasks. *Journal of Speech, Language, and Hearing Research*. 2005; 48:295–305.
- Dromey C, Benson A. Effects of concurrent motor, linguistic, or cognitive tasks on speech motor performance. *Journal of Speech, Language, and Hearing Research*. 2003; 46:1234–1246.
- Dromey C, Ramig LO. Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation, and articulation. *Journal of Speech, Language, and Hearing Research*. 1998; 41:1003–1018.
- Dromey C, Ramig LO, Johnson AB. Phonatory and articulatory changes associated with increased vocal intensity in Parkinson's Disease: A case study. *Journal of Speech and Hearing Research*. 1995; 38:751–764. [PubMed: 7474969]
- Fant, G. *Acoustic Theory of Speech Production*. 2nd ed.. Paris: Mouton; 1970.
- Forrest K, Weismer G, Turner GS. Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric speakers. *The Journal of the Acoustical Society of America*. 1989; 85(6):2608–2622. [PubMed: 2745883]
- Gentilucci M, Benuzzi F, Bertolani L, Daprati E, Gangitano M. Language and motor control. *Experimental Brain Research*. 2000; 133:468–490.
- Green JR, Moore CA, Higashikawa M, Steeve RW. The physiologic development of speech motor control: Lip and jaw coordination. *Journal of Speech, Language, and Hearing Research*. 2000; 43:239–255.

- Huber JE, Chandrasekaran B, Wolstencroft JJ. Changes to respiratory mechanisms during speech as a result of different cues to increase loudness. *Journal of Applied Physiology*. 2005; 98:2177–2184. [PubMed: 15705723]
- Huber JE, Stathopoulos ET, Curione GM, Ash TA, Johnson K. Formants of children, women, and men: The effects of vocal intensity variation. *The Journal of the Acoustical Society of America*. 1999; 106(3, Pt. 1):1532–1542. [PubMed: 10489709]
- Kleinow J, Smith A, Ramig LO. Speech motor stability in IPD: Effects of rate and loudness manipulations. *Journal of Speech, Language, and Hearing Research*. 2001; 44:1041–1051.
- Lane H, Tranel B. The Lombard Sign and the role of hearing in speech. *Journal of Speech and Hearing Research*. 1971; 14:677–709.
- McClellan MD, Tasko SM. Association of orofacial muscle activity and movement during changes in speech rate and intensity. *Journal of Speech, Language, and Hearing Research*. 2003; 46:1387–1400.
- Mendel LL, Walton JH, Hamill BW, Pelton JD. Use of speech production repair strategies to improve diver communication. *Undersea & Hyperbaric Medicine*. 2003; 30:313–320. [PubMed: 14756234]
- Parkhurst BG, Levitt H. The effect of selected prosodic errors on the intelligibility of deaf speech. *Journal of Communication Disorders*. 1978; 11:249–256. [PubMed: 659657]
- Perkell JS, Matthies ML, Svirsky MA, Jordan MI. Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot "motor equivalence" study. *The Journal of the Acoustical Society of America*. 1993; 93(5):2948–2961. [PubMed: 8315158]
- Pick HL Jr, Siegel GM, Fox PW, Garber SR, Kearney JK. Inhibiting the Lombard effect. *The Journal of the Acoustical Society of America*. 1989; 85(2):894–900. [PubMed: 2926004]
- Ramig LO, Countryman S, Thompson LL, Horii Y. Comparison of two forms of intensive speech treatment for Parkinson Disease. *Journal of Speech and Hearing Research*. 1995; 38:1232–1251. [PubMed: 8747817]
- Schulman R. Articulatory dynamics of loud and normal speech. *The Journal of the Acoustical Society of America*. 1989; 85:295–312. [PubMed: 2921412]
- Smith A, Goffman L, Zelaznik HN, Ying G, McGillen CI. Spatiotemporal stability and patterning of speech movement sequences. *Experimental Brain Research*. 1995; 104:493–501.
- Smith A, Johnson M, McGillem C, Goffman L. On the assessment of stability and patterning of speech movements. *Journal of Speech, Language, and Hearing Research*. 2000; 43:277–286.
- Stathopoulos ET, Sapienza CM. Developmental changes in laryngeal and respiratory function with variations in sound pressure level. *Journal of Speech, Language, and Hearing Research*. 1997; 40:595–614.
- Stevens SS. The measurement of loudness. *The Journal of the Acoustical Society of America*. 1955; 27:815–829.
- Tasko SM, McClellan MD. Variations in articulatory movement with changes in speech task. *Journal of Speech, Language, and Hearing Research*. 2004; 47:85–100.
- Van Summers W, Pisoni DB, Bernacki RH, Pedlow RI, Stokes MA. Effects of noise on speech production: Acoustic and perceptual analyses. *The Journal of the Acoustical Society of America*. 1988; 84(3):917–928. [PubMed: 3183209]
- Walsh B, Smith A. Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research*. 2002; 45:1119–1133.
- Winkworth AL, Davis PJ. Speech breathing and the Lombard effect. *Journal of Speech, Language, and Hearing Research*. 1997; 40:159–169.
- Wohlert AB, Smith A. Spatiotemporal stability of lip movements in older adult speakers. *Journal of Speech, Language, and Hearing Research*. 1998; 41:41–50.
- Yorkston KM, Hammen VL, Beukelman DR, Traynor C. The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders*. 1990; 55:550–560. [PubMed: 2381196]

SHORT sentence:



EMBED sentence:

**Figure 1.**

Segmentation and measurement points for the lip and jaw kinematic traces. Trace of lower lip + jaw from female speaker during comfortable condition. Top: SHORT sentence; Bottom: EMBED sentence. Lines A and B are the segmentation points. Point C is peak closure for the first “b” in “Bob”; point D is the peak opening for the vowel in “Bob”; point E is the peak closure for the second “b” in “Bob.” Point F is peak opening velocity for the vowel in “Bob”; point G is peak closing velocity for the second “b” in “Bob.”

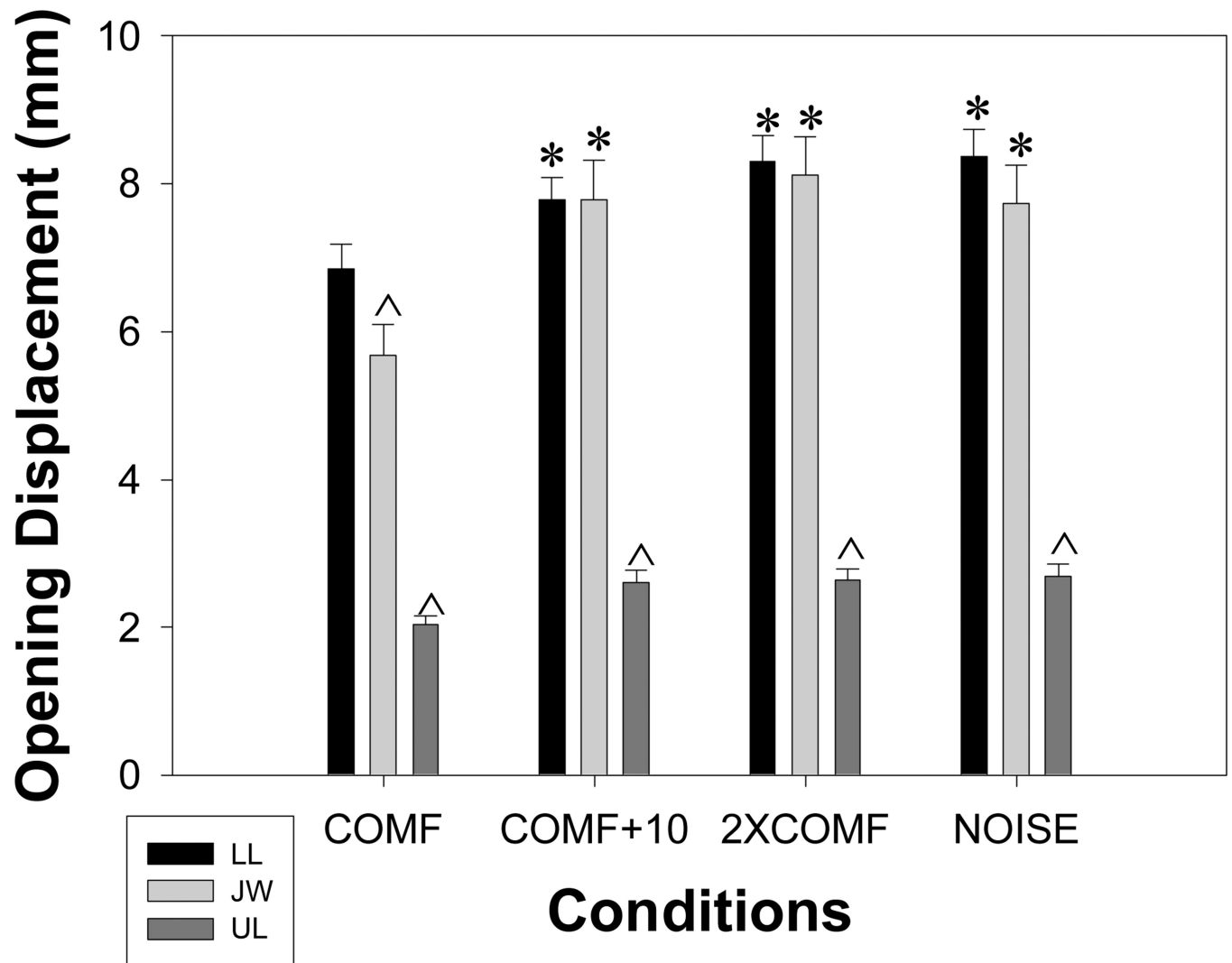


Figure 2. Means for opening displacement (ODISP) articulator by condition interaction, LL = lower lip, JW = jaw, and UL = upper lip. Bars represent means; lines show standard errors. Asterisks indicate significantly different from COMF within articulator; carot symbols indicate significantly different from lower lip within condition.

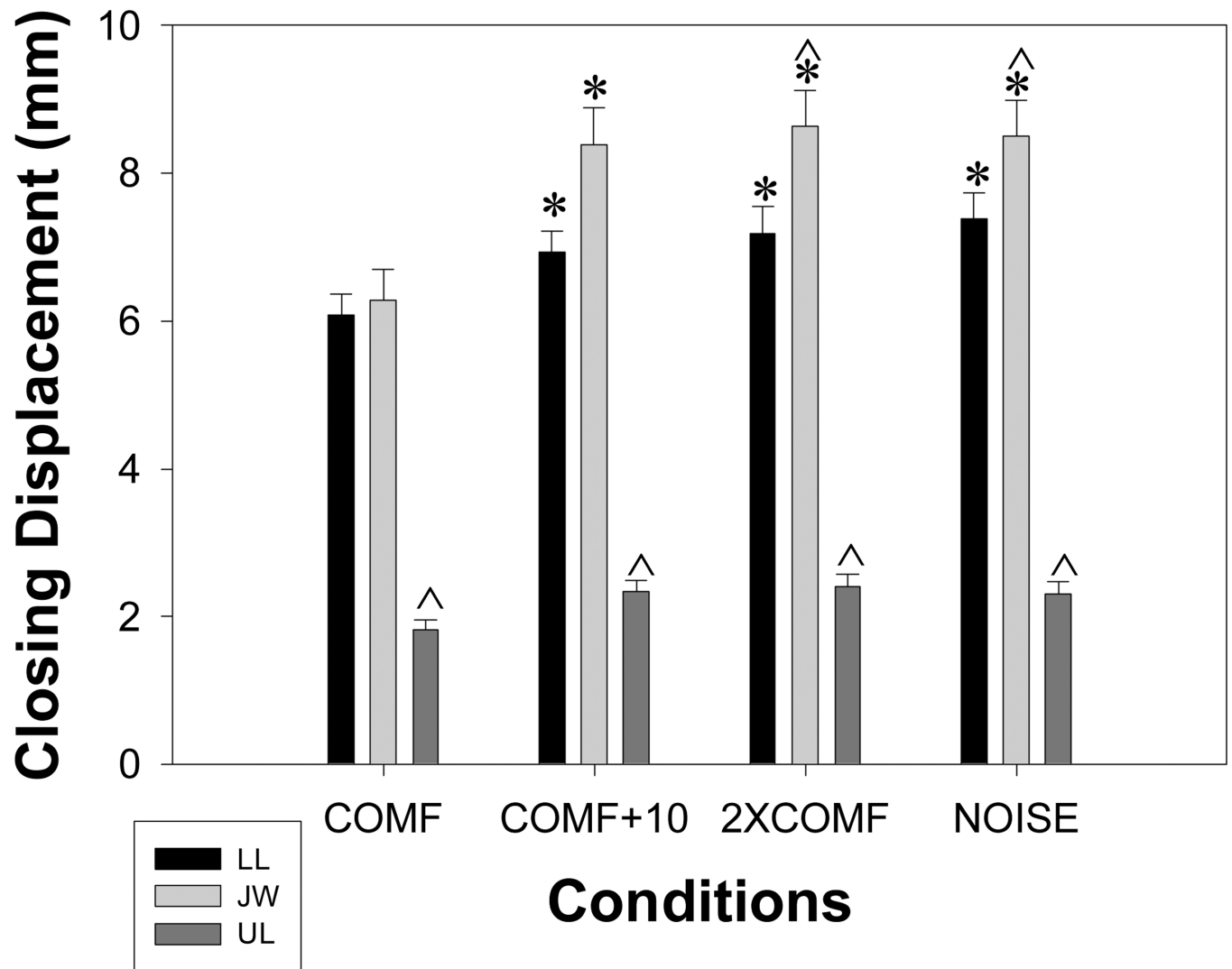


Figure 3. Means for the closing displacement (CDISP) articulator by condition interaction, LL = lower lip, JW = jaw, and UL = upper lip. Bars represent means; lines show standard errors. Asterisks indicate significantly different from COMF within articulator; carot symbols indicate significantly different from lower lip within condition.

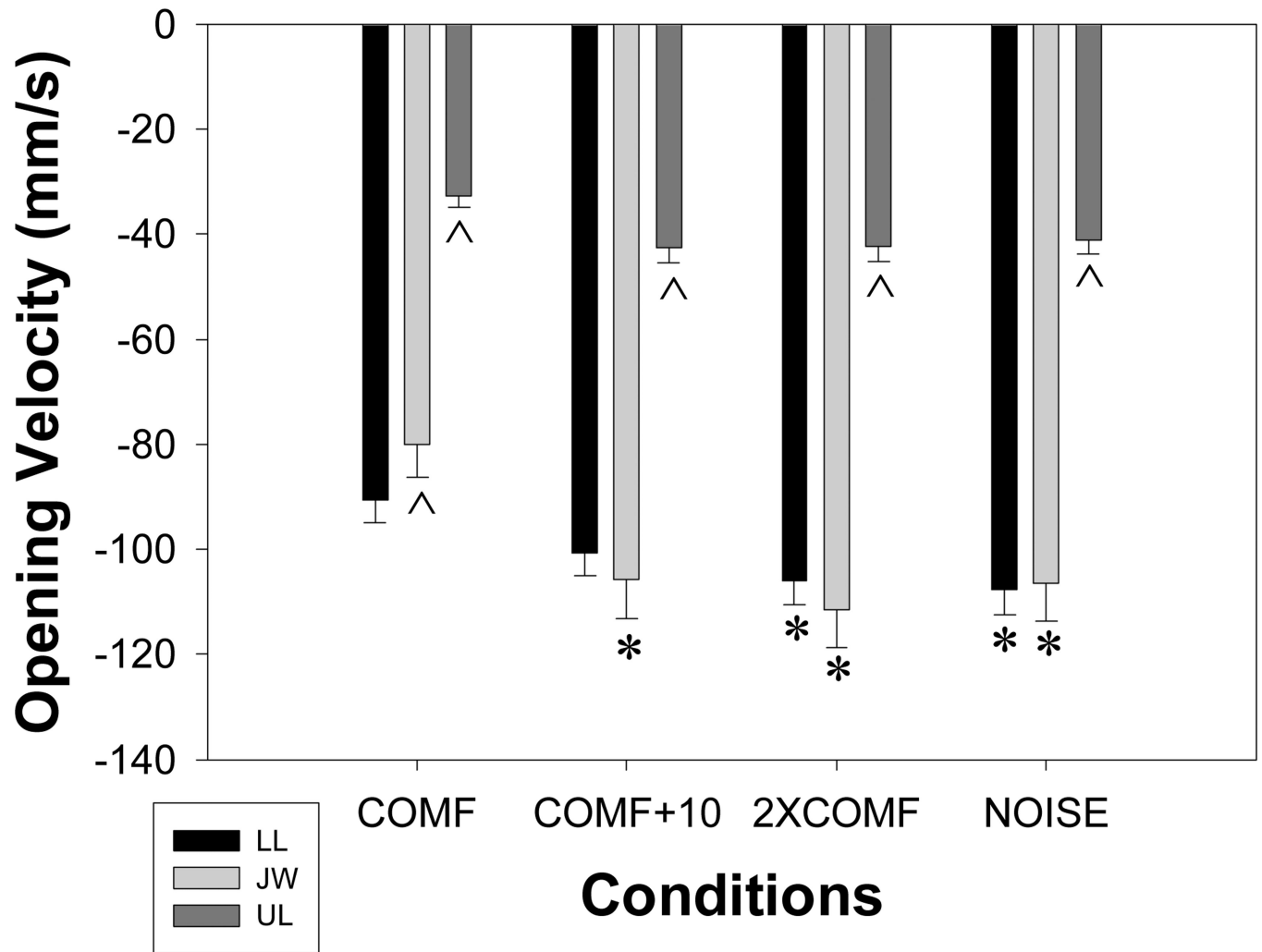


Figure 4. Means for the opening velocity (OVEL) articulator by condition interaction, LL = lower lip, JW = jaw, and UL = upper lip. Bars represent means; lines show standard errors. Asterisks indicate significantly different from COMF within articulator; carots indicate significantly different from lower lip within condition.

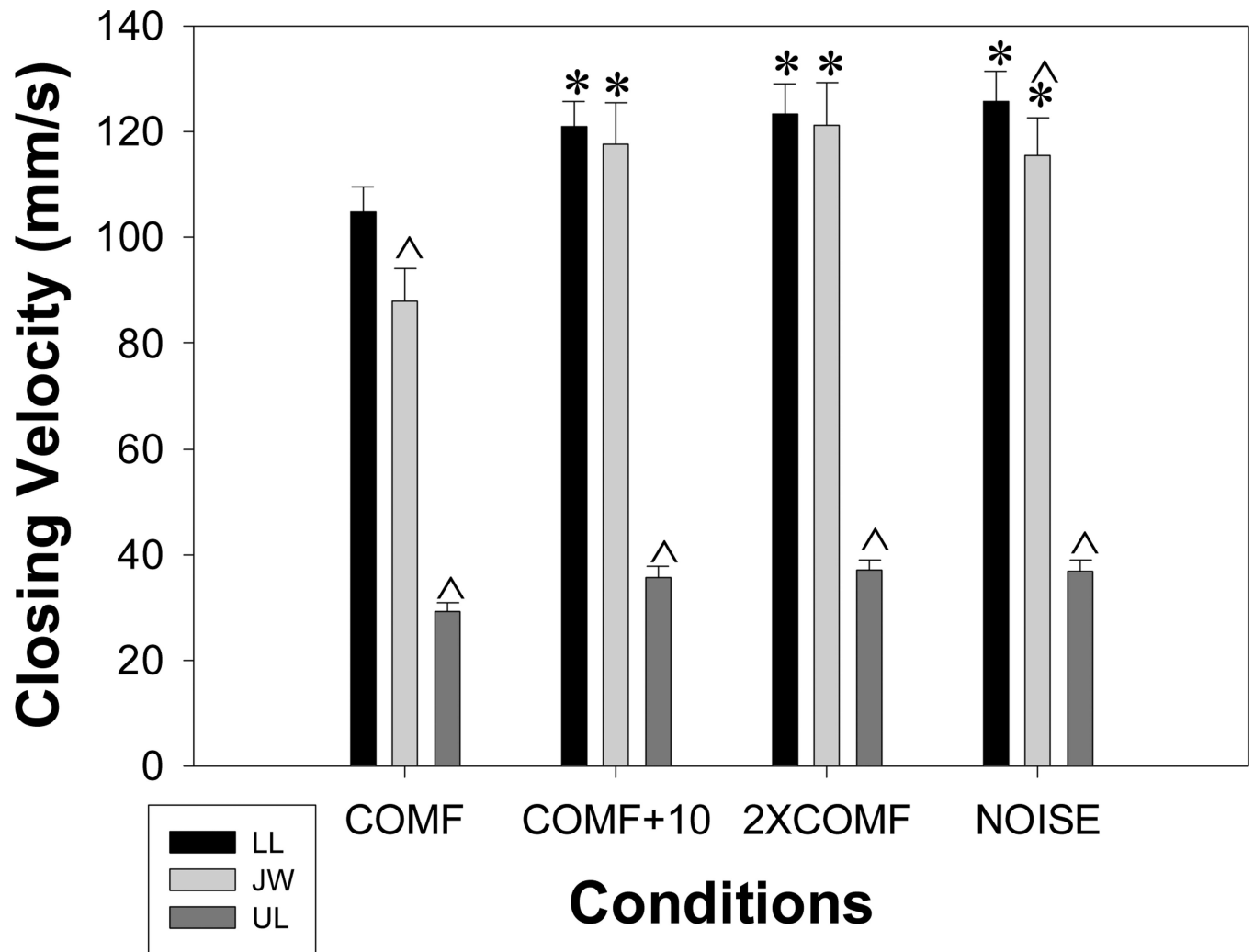


Figure 5.

Means for the closing velocity (CVEL) articulator by condition interaction, LL = lower lip, JW = jaw, and UL = upper lip. Bars represent means; lines show standard errors. Asterisks indicate significantly different from COMF within articulator; carot symbols indicate significantly different from lower lip within condition.

Table 1

Means and standard deviations (in parentheses) for conditions.

Measurements	COMF	COMF+10	2XCOMF	NOISE
Sound Pressure Level (dB)	79.09 (3.98)	89.32* (4.59)	88.44* (5.50)	90.20* (5.44)
First Formant Frequency (Hz)	739.99 (110.42)	817.04* (127.17)	818.66* (137.18)	834.31* (127.17)
Second Formant Frequency (Hz)	1231.55 (162.57)	1310.47* (189.60)	1299.76 (159.46)	1300.12 (189.60)
Sentence Duration (s)	1.75 (0.25)	1.82 (0.32)	1.78 (0.30)	1.85* (0.32)
“Bob” Duration (s)	0.22 (0.04)	0.23 (0.04)	0.23 (0.04)	0.23 (0.04)
Opening Displacement (mm)	4.86 (3.20)	6.06* (3.80)	6.35* (3.96)	6.27* (3.89)
Closing Displacement (mm)	4.73 (3.14)	5.88* (3.72)	6.08* (3.93)	6.07* (3.85)
Opening Velocity (mm/s)	-67.92 (43.16)	-83.04* (50.53)	-86.72* (51.40)	-85.17* (51.84)
Closing Velocity (mm/s)	74.09 (48.71)	91.41* (58.09)	93.94* (60.58)	92.75* (58.33)
Spatiotemporal Index (STI)	17.91 (5.76)	15.61* (6.31)	16.66 (6.31)	17.01 (5.77)

COMF = comfortable loudness condition; COMF+10 = ten decibels above comfortable loudness; 2XCOMF = twice as loud as comfortable loudness; NOISE = in noise; dB = decibels; Hz = Hertz; s = seconds; mm = millimeters.

* significantly different from COMF

Table 2

Means and standard deviations (in parentheses) for articulators.

Measurements	Lower Lip	Jaw	Upper Lip
"Bob" Duration (s)	0.22 (0.04) [*]	0.25 (0.04) ^{*+}	0.21 (0.04) ⁺
Opening Displacement (mm)	7.84 (2.81) [*]	7.32 (3.97) [*]	2.49 (1.23)
Closing Displacement (mm)	6.91 (2.65) [*]	7.95 (3.76) [*]	2.22 (1.23)
Opening Velocity (mm/s)	-101.45 [*] (37.35)	-100.95 [*] (55.66)	-39.74 (21.13)
Closing Velocity (mm/s)	118.89 [*] (43.08)	110.54 [*] (58.21)	34.72 (15.37)
Spatiotemporal Index (STI)	15.83 [*] (4.48)	13.56 [*] (3.76)	21.01 (6.91)

dB = decibels; Hz = Hertz; s = seconds; mm = millimeters.

^{*} significantly different from upper lip⁺ significantly different from lower lip

Table 3

Statistical Summary

Measurements	Condition (3, 87) a (3, 180) k		Articulator (2, 120) k		Condition X Articulator (6, 360) k	
	F	p	F	p	F	p
Sound Pressure Level	124.6	0.00*	n/a	n/a	n/a	n/a
First Formant Frequency	25.7	0.00*	n/a	n/a	n/a	n/a
Second Formant Frequency	4.6	0.00*	n/a	n/a	n/a	n/a
Sentence Duration	4.6	0.01*	n/a	n/a	n/a	n/a
“Bob” Duration	3.0	0.03	87.2	0.00*	1.0	0.42
Opening Displacement	37.7	0.00*	84.2	0.00*	10.2	0.00*
Closing Displacement	37.0	0.00*	102.1	0.00*	11.9	0.00*
Opening Velocity	34.7	0.00*	66.5	0.00*	7.6	0.00*
Closing Velocity	40.3	0.00*	95.3	0.00*	10.2	0.00*
Spatiotemporal Index (STI)	5.8	0.00*	26.9	0.00*	2.4	0.03

Degrees of freedom in parentheses, “a” for acoustic measurements, “k” for kinematic measurements. F= F-ratio, p= level of significance. Asterisk indicates significance at the $p = 0.01$ level.