Changes to respiratory mechanisms during speech as a result of different cues to increase loudness

Jessica E. Huber, Bharath Chandrasekaran, and John J. Wolstencroft
Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, Indiana

Submitted 4 November 2004; accepted in final form 8 February 2005

Huber, Jessica E., Bharath Chandrasekaran, and John J. Wolstencroft. Changes to respiratory mechanisms during speech as a result of different cues to increase loudness. J Appl Physiol 98: 2177–2184, 2005. First published February 10, 2005; doi:10.1152/japplphysiol.01239.2004.—The purpose of the present study was to determine whether different cues to increase loudness in speech result in different internal targets (or goals) for respiratory movement and whether the neural control of the respiratory system is sensitive to changes in the speaker’s internal loudness target. This study examined respiratory mechanisms during speech in 30 young adults at comfortable level and increased loudness levels. Increased loudness was elicited using three methods: asking subjects to target a specific sound pressure level, asking subjects to speak twice as loud as comfortable, and asking subjects to speak in noise. All three loud conditions resulted in similar increases in sound pressure level. However, the respiratory mechanisms used to support the increase in loudness differed significantly depending on how the louder speech was elicited. When asked to target at a particular sound pressure level, subjects used a mechanism of increasing the lung volume at which speech was initiated to take advantage of higher recoil pressures. When asked to speak twice as loud as comfortable, subjects increased expiratory muscle tension, for the most part, to increase the pressure for speech. However, in the most natural of the elicitation methods, speaking in noise, the subjects used a combined respiratory approach, using both increased recoil pressures and increased expiratory muscle tension. In noise, an additional target, possibly improving intelligibility of speech, was reflected in the slowing of speech rate and in larger volume excursions even though the speakers were producing the same number of syllables.

The examination of how the respiratory system supports speech production is a long-standing area of study. Previous studies have demonstrated that the respiratory system provides a steady driving pressure for speech production by balancing recoil and muscular pressures in the system (3, 8, 9, 13, 19). Loud speech requires a higher subglottic pressure. To achieve this, most individuals begin to speak at a higher lung volume (LV) to take advantage of higher recoil pressures (4, 8, 19), although all speakers do not increase their LV to the same degree (19). A second mechanism, more often used by speakers closer to or below end-expiratory level (EEL), is to increase expiratory muscle tension to achieve the required higher driving pressures.

However, not all studies have demonstrated such clear trends in respiratory function for loud speech. For example, Winkworth and Davis (20) reported variable respiratory patterns in five women speaking in noise. Speaking in noise is known to elicit the Lombard effect in which speakers naturally speak louder under conditions of background noise (12, 18). The variability in the respiratory patterns in the Winkworth and Davis study was distinctly different from previous studies that reported a pattern of initiating speech at higher LVs. The subjects in the Winkworth and Davis study increased their sound pressure level (SPL) by 10–16 dB, which was at least as much as the increases in SPL reported in previous studies of respiratory patterns.

Winkworth and Davis (20) hypothesized that the differences between their findings and earlier findings were due to the method used to elicit louder speech. The bulk of the data on respiratory kinematics and increasing loudness have used a small set of cues: 1) asking subjects to speak twice or four times as loud as comfortable (4) and 2) asking individuals to target a specific SPL (usually 5–10 dB higher than their comfortable level) on an SPL meter (19). Cues to speak twice or four times as loud have resulted in increased LV initiations (LVIs) (4), and targeting cues have resulted in increased volume initiations and larger volume excursions (19).

The type of cue used to elicit an increase in loudness has not been studied previously mainly because of the assertion that respiratory kinematics do not change as a result of how an individual is cued to increase loudness. The findings from Winkworth and Davis (20) suggest that this may not be true. Their data demonstrate that multiple strategies may be employed by speakers to speak louder when they are not specifically instructed to increase their loudness.

There are many potential differences between instructing a subject to speak louder and placing a subject in an environment that automatically triggers an increase in loudness. Instructing subjects to speak louder may cause them to focus on being loud more so than they would in a natural situation, such as when louder speech is required to avoid communicative breakdown (in a noisy environment). Speaking in noise may involve communication goals not elicited by an instruction to increase loudness. These additional goals may include increasing the intelligibility of the speech signal to improve its transmission through noise (12). Finally, speakers may perceive the difficulty of the goal of increasing loudness differently depending on the situation. Being asked to increase loudness may be perceived as more difficult than speaking in noise, in which the increase in loudness is viewed as more automatic (18). Additionally, speaking at 10 dB above comfortable (typically 85–90 dB SPL) may sound harder to subjects than speaking twice as loud as their comfortable loudness level.

Any of the above differences could potentially change the speaker’s internal target. An internal target is the neural representation of the desired movement parameters and reflects the goals of the individual’s movements. In the present study, the costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
the goal is to be louder. The internal target results in a scaling up of the movement plans that go out to the muscles in the periphery. It is possible that the way an individual is cued may affect internal targets and, therefore, how the movement parameters are modified. For example, in a noisy environment, two targets may exist: to increase loudness and to slow speech to improve transmission of the message. In this case, the internal target may be for louder and slower speech, and movement parameters would be modified to be both scaled up and slowed down.

Studies of limb movement have demonstrated that changes to peripheral movements provide information about internal targets (7). For example, in the study by Gentilucci and colleagues (7), reaching and grasping movement was affected by the label on the object, such as “large” or “small,” regardless of the actual size of the object. In this case, automatic word reading had a greater effect on the internal target than the actual perception of the object, and the internal target was reflected in the movement parameters for reaching and grasping. Because the internal target of a movement is reflected in the kinematics associated with achieving that target, it is possible that the movements of the respiratory system will be altered based on changes to the internal target associated with specific cues to increase loudness.

The purpose of the present study was to determine whether different cues to increase loudness will result in different respiratory kinematic patterns. Examination of mechanisms for increasing loudness under different cues is important from a motor control perspective since a greater understanding of the respiratory system’s sensitivity to internal targets will enhance our knowledge of its role in speech production. Unique kinematic patterns, based on cue, would support the view that cues create different internal targets and that the neural control of the respiratory system is sensitive to changes in internal targets for speech.

The aim of the present study is also important from a clinical perspective in speech therapy. There are a number of speech disorders that result in reduced loudness levels, such as Parkinson’s disease. Individuals with symptoms of low vocal loudness often undergo therapies that cue them, in many ways, to increase their loudness. Because the respiratory system plays a primary role in increasing loudness, an understanding of how cues affect respiratory function in normal speakers will assist in choosing the best cues for treatment.

MATERIALS AND METHODS

Subjects

Thirty normal young adults, 15 women and 15 men, participated in the study. The mean age of the women was 22 yr, 4 mo, and the mean age of the men was 22 yr, 10 mo. Speakers were grouped by sex because previous studies have demonstrated differences in respiratory kinematics during speech between the sexes (10, 19).

Subjects indicated that they had 1) no history of voice or respiratory problems (including asthma), neurological disease, or head or neck surgery; 2) never received formal speaking or singing training; 3) no recent colds or infections; and 4) been nonsmoking for the past 5 yr. They had a body mass index between 19 and 30 as measured on the day of testing (5), normal speech, language, and voice, and spoke General North American dialect of English. They had normal hearing as indicated by a hearing screening at 30 dB hearing level for octave frequencies between 250 and 8,000 Hz, bilaterally, completed in a quiet room. Subjects had normal vital capacity (VC), forced VC, and forced expiratory volume in 1 s defined as ≥80% of expected values based on age, gender, height, weight, and ethnicity. Lung capacities were measured using a digital spirometer (SP; VacuMed Discovery Handheld SP).

Procedures and Speech Stimuli

Procedures for data collection were approved by Purdue University’s Committee on the Use of Human Research Subjects. Subjects said two sentences: 1) “Buy Bobby a puppy” and 2) “You buy Bobby a puppy now if he wants one.” Use of two sentences of specified length ensured that utterance length was controlled and was not a factor in the kinematic measurements. Subjects were instructed to say one sentence per breath and to speak clearly and audibly in each condition. The experimenter was visible to the subject during all conditions and was seated about 80 in. away. The conditions were as follows.

Comfortable level. Subjects were instructed to “read the sentence at your comfortable loudness and pitch” (COMF).

COMF + 10. Subjects 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 for the XX in the instruction. The SPL targets for this condition were set at 10 dB (+2 dB) above the subject’s comfortable SPL. The SPL meter was set at fast response. The output from the SPL meter was enlarged and projected onto a television screen that faced the subject. Feedback was provided continuously to the subject.

2× COMF. Subjects were instructed to “read the sentence at what you feel is twice your comfortable loudness.” No feedback was provided during this condition. If this condition did not follow the COMF condition, subjects were instructed to “read the sentence at your comfortable loudness and pitch” until their SPL level was similar to their original SPL level for the COMF condition before being given the twice-as-loud cue. The decision of similarity between the SPL levels was made by the examiner using the SPL meter.

NOISE. Multi-talker noise (AUDITEC, St. Louis, MO) was turned on in the room. The noise was delivered at 70 dBA relative to the subject’s ears via free field. Subjects were instructed to “read the sentence.” No cue was given regarding loudness in this condition. The speakers used to deliver the noise were placed 39 in. in front of the subjects.

For each condition, the shorter sentence was completed first, and each sentence was said 15 times consecutively. The COMF condition was always completed first. The order of the three loud conditions (COMF + 10, 2× COMF, and NOISE) was counterbalanced across subjects.

Subjects also produced maximum capacity tasks (9). These tasks were used to obtain an estimate of the maximal capacity of the lungs, rib cage (RC), and abdomen (Ab). Measurements during speech production were expressed as a percent of capacity so that comparisons could be made across individuals of differing sizes. In all cases, subjects were expected to produce three comparable maximum capacity tasks. VC measured as a part of the inclusion criteria was used, and LV measures were expressed as a percentage of the largest VC produced. VC maneuvers were also completed at least three times with the Respitrace bands in place. These trials were in addition to those used to obtain VC for subject inclusion and were used to obtain RC capacity (RCC) since the RC moves maximally during a VC maneuver. To determine abdominal (Ab) capacity (AbC), two maneuvers were completed, at least three times each, with the Respitrace bands in place. First, the subject was instructed to hold his/her breath at EEL and suck his/her stomach in maximally. Second, the subject was instructed to hold his/her breath at EEL and extend his/her stomach maximally. The combination of the maximum in and the maximum out were taken to be total AbC. At least three steady cycles
of rest breathing (RB) were collected before the start of each trial of the maximum capacity maneuvers.

**Equipment**

The acoustic signal was transduced via a condenser microphone that was connected to an SPL meter (Quest model 1700). The microphone was placed 6 in. from the subject’s mouth at a 45° angle. The microphone signal was recorded to digital audiotape and later digitized into a personal computer using Praat (1). The signal was digitized at 44.1 kHz and resampled at 18 kHz. The resampling process applied a low-pass filter at 9,000 Hz for antialiasing.

Respiratory kinematic data were transduced via respiratory inductive plethysmography using the Respirtrace system (Ambulatory Monitoring). An elastic band was placed around the RC, just under the axilla, to transduce movements of the RC. A second elastic band was placed around the abdomen at the level of the umbilicus, ensuring that it was below the last rib, to transduce movements of the abdomen. Respiratory kinematic data were digitized at 2,000 Hz. Data from a second microphone were collected with these data so that an acoustic record would be digitized in combination with the respiratory kinematic data.

**Measurements**

The first two trials of each sentence in each condition were discarded. The next 10 consecutive sentences that were produced without error were chosen for analysis. SPL was measured using Praat (1) for each sentence.

**Respiratory kinematic measurements.** Respiratory kinematic measurements were made using algorithms written in Matlab (MathWorks). Before any measurements were made, the respiratory kinematic signals were low-pass filtered at 40 Hz to remove noise.

Because LV change reflects the combined effect of changes in the RC and Ab volumes (11), the sum of the RC and Ab signals was computed and corrected for the respective RC and Ab contributions to LV change. Two nonspeech tasks were used to determine the RC and Ab contributions to LV change. Subjects were instructed to relax, and data were collected for two 45-s periods of RB. Data was also collected for three 45-s periods of “speech-like” breathing (SLB). For this task, subjects were instructed to read the longer sentence silently to themselves, one time per breath. At least three steady cycles of RB were collected before the start of each SLB data collection period. LV data was collected during the RB and SLB tasks using a SP (VacuMed Universal Ventilation Meter). This SP has a very small dead space. These data were digitized along with the respiratory kinematic data at 2,000 Hz.

The data from the SP, RC, and Ab signals during the RB and SLB tasks were used to determine the correction factors for the RC and Ab. The Moore-Penrose pseudoinverse function was used in Matlab to determine the least errored solution for the correction factors (k$_1$ and k$_2$). The pseudoinverse function solved for k$_1$ and k$_2$ in the formula SP = k$_1$(RC) + k$_2$(Ab) for each set of RC, Ab, and SP data points in the RB and SLB tasks. This estimation of the LV signal was verified by visually checking the SUM signal [k$_1$(RC) + k$_2$(Ab)] against the original SP signal for a SLB trial. The estimated LV signal was then computed for each point during the sentence production tasks using the formula LV = k$_1$(RC) + k$_2$(Ab).

LV, RC, and Ab inspirations and terminations were measured relative to EEL. EEL was measured from troughs of three steady RB before the start of each set of sentence repetitions. LV, RC, and Ab initiations and terminations were expressed as a percent of VC, RCC, and AbC, respectively. Speech initiations were defined as the point where voicing began, and speech terminations were defined as the point where voicing ended, as indicated by the microphone signal collected with the kinematic data (see Fig. 1, lines A and B). The audio signal was used to verify that the initiations and terminations were selected accurately and that no part of the speech signal was cut off.

Fig. 1. Measurement points for the respiratory kinetic traces. Trace from speaker during the comfortable condition, longer sentence. Top: line A is the point where lung (LV), rib cage (RC), and abdominal (Ab) volume initiations were measured. Line B is the point where LV, RC, and Ab volume terminations were measured. Bottom: point C (on LV waveform) is the volume at the end of expiration for the previous utterance. Point D is the top of inspiration for the current utterance.

LV, RC, and Ab excursions were calculated as the volume at initiation minus the volume at termination and expressed as a percent of VC, RCC, and AbC, respectively (see Fig. 1, line A to line B).

Percent VC expended per syllable was measured by dividing the LV excursion by the number of syllables for each utterance. Percent VC inspired was measured from the LV signal by subtracting the volume at the end of expiration after the previous utterance from the volume at the top of inspiration before the current utterance (see Fig. 1, point D to point C).

**Timing Measurements**

Duration was measured as the time between speech initiation and speech termination of each utterance. Syllables per second was measured by dividing the duration of the utterance by the number of syllables produced. The phonation onset time was defined as the time from the end of inspiration to the start of speech for the sentence (see Fig. 1, point D to line A).

**Statistics**

Means were computed for each subject for each condition. The differences in the means were assessed in two-factor repeated-measures ANOVA. The within factor was loudness (condition) and the between factor was sex. Table 1 provides a summary of the statistical results. Tukey’s honestly significant difference tests were completed for all factors and interactions that were significant in the ANOVA. The alpha level for the ANOVAs and the Tukey’s honestly significant difference tests were set at $P \leq 0.01$.

To test for a learning effect across the 10 trials used for analysis, a matched-pairs t-test was computed between the mean of the first three trials and the mean of the last three trials for each measurement. The alpha level was set at $P \leq 0.01$, as it was for the ANOVAs. There were no significant differences between the two means for any of the measurements, suggesting that there was no significant learning effect.

Intermeasurer reliability was completed on two male and two female subjects, randomly chosen. Independent t-tests were computed between the first and second measurement for each variable. None of the alpha levels neared significance, ranging from $P = 0.128$ to $P = 0.881$, indicating good intermeasurer reliability.
RESULTS

For SPL, there was a significant condition effect but no significant sex or interaction effects. The three loud conditions were produced at a significantly higher SPL than the COMF condition. There were no significant differences in SPL for the three loud conditions. The mean SPLs for the conditions were 79 dB (SE = 0.55 dB) for COMF, 89 dB (SE = 0.61 dB) for COMF +10, 88 dB (SE = 0.73 dB) for 2× COMF, and 90 dB (SE = 0.71 dB) for NOISE.

Timing Measurements

For duration, there was a significant condition effect but no significant sex or interaction effects. The sentences produced in the NOISE condition were produced over a significantly longer duration than those in the COMF condition (see Fig. 2). For syllables per second, there was a significant condition effect but no sex or interaction effects. Significantly fewer syllables per second were produced in the NOISE condition compared with the COMF condition (see Fig. 2). For phonation onset time, there was a significant condition effect but no sex or interaction effects. The phonation onset time was significantly shorter in the COMF +10 and 2× COMF conditions compared with the COMF condition (see Fig. 2).

Respiratory Kinematic Measurements

For LVI, there were significant condition and sex effects but no interaction effect. LVI was significantly higher in the COMF +10 condition compared with the COMF condition (see Fig. 3). LVI was significantly higher for the women compared with the men.

For LV termination (LVT), there were significant condition and sex effects but no interaction effect. LVT was significantly higher for the COMF +10 condition compared with the COMF condition (see Fig. 3). LVT was significantly higher for the women than for the men.

For LV excursion (LVE), there was a significant condition effect but no sex or interaction effects. LVE was significantly larger in the NOISE condition than in the COMF condition (see Fig. 4).

For percent of VC inspired, there was a significant condition effect but no sex or interaction effects. The percent of VC inspired was significantly larger in the NOISE condition than in the COMF and COMF +10 conditions (see Fig. 5).

For percent VC expended per syllable, there was a significant condition effect but no sex or interaction effects. Percent

Table 1. Statistical summary for condition, sex, and condition by sex interactions

<table>
<thead>
<tr>
<th>Measures</th>
<th>Condition (3, 84)</th>
<th>Sex (1, 28)</th>
<th>Condition × Sex (3, 84)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>P</td>
<td>F</td>
</tr>
<tr>
<td>Sound pressure level</td>
<td>122.51</td>
<td>0.000*</td>
<td>0.27</td>
</tr>
<tr>
<td>Duration</td>
<td>4.57</td>
<td>0.005*</td>
<td>2.76</td>
</tr>
<tr>
<td>Syllables per second</td>
<td>4.63</td>
<td>0.005*</td>
<td>3.81</td>
</tr>
<tr>
<td>Phonation onset time</td>
<td>5.34</td>
<td>0.002*</td>
<td>0.22</td>
</tr>
<tr>
<td>Lung volume initiation</td>
<td>6.65</td>
<td>0.000*</td>
<td>17.10</td>
</tr>
<tr>
<td>Lung volume termination</td>
<td>5.00</td>
<td>0.003*</td>
<td>14.73</td>
</tr>
<tr>
<td>Lung volume excursion</td>
<td>7.87</td>
<td>0.000*</td>
<td>3.82</td>
</tr>
<tr>
<td>Percent vital capacity inspired</td>
<td>9.30</td>
<td>0.000*</td>
<td>2.60</td>
</tr>
<tr>
<td>Percent vital capacity</td>
<td>8.89</td>
<td>0.000*</td>
<td>3.23</td>
</tr>
<tr>
<td>expended per syllable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rib cage volume initiation</td>
<td>8.17</td>
<td>0.000*</td>
<td>5.31</td>
</tr>
<tr>
<td>Rib cage volume termination</td>
<td>4.25</td>
<td>0.008*</td>
<td>5.96</td>
</tr>
<tr>
<td>Rib cage volume excursion</td>
<td>12.51</td>
<td>0.000*</td>
<td>0.14</td>
</tr>
<tr>
<td>Abdominal volume initiation</td>
<td>6.44</td>
<td>0.001*</td>
<td>0.20</td>
</tr>
<tr>
<td>Abdominal volume termination</td>
<td>5.29</td>
<td>0.002*</td>
<td>0.06</td>
</tr>
<tr>
<td>Abdominal volume excursion</td>
<td>0.44</td>
<td>0.723</td>
<td>1.88</td>
</tr>
</tbody>
</table>

Degrees of freedom shown in parentheses. F, F ratio; P, level of significance. *Significance at the P ≤ 0.01 level.

Fig. 2. Timing measurements: mean difference from comfortable speaking level (COMF) for the 3 loud conditions (COMF +10 dB (COMF +10), twice the COMF volume (2× COMF), and speaking over background noise (NOISE)). Lines show standard errors. *Significant changes from COMF (P ≤ 0.01).

Fig. 3. Initiation and termination measurements: mean difference from COMF for the 3 loud conditions (COMF +10, 2× COMF, and NOISE). Lines show standard errors. LVI, lung volume initiation; LVT, lung volume termination; RCVI, rib cage volume initiation; RCVT, rib cage volume termination; ABVI, abdominal volume initiation; ABVT, abdominal volume termination. *Significant changes from COMF (P ≤ 0.01).
VC expended per syllable was significantly higher in the NOISE condition compared with the COMF condition (see Fig. 5).

For RC volume initiation (RCVI), there was a significant condition effect but no sex or interaction effects. RCVI was significantly higher in the COMF + 10 and NOISE conditions compared with COMF condition (see Fig. 3).

For RC volume termination (RCVT), there was a significant condition effect but no sex or interaction effects. RCVT was significantly higher in the COMF + 10 condition compared with the COMF condition (see Fig. 3).

For RC volume excursion (RCVE), there was a significant condition effect but no sex or interaction effects. RCVE was significantly larger in all loud conditions compared with the COMF condition (see Fig. 4).

For Ab volume initiation (AbVI), there was a significant condition effect but no sex or interaction effects. AbVI was significantly lower in the 2× COMF condition compared with the COMF condition (see Fig. 3).

For Ab volume termination (AbVT), there was a significant condition effect but no sex or interaction effects. AbVT was significantly lower in the 2× COMF condition compared with the COMF condition (see Fig. 3).

For Ab volume excursion, there were no condition, sex, or interaction effects (see Fig. 4).

**DISCUSSION**

The purpose of the present study was to test whether the kinematic patterns of the chest wall will differ based on how speakers were cued to increase loudness. All three loud conditions resulted in similar SPL increases: ~10 dB. However, the respiratory mechanisms used to increase loudness differed depending how the increase in loudness was elicited. The different kinematic patterns suggest that the cues resulted in different internal targets and that the neural control of the respiratory system for speech is affected by changes to the speaker’s internal loudness target.

In the COMF + 10 condition, subjects primarily used a mechanism of beginning to speak at a higher LV to utilize higher recoil pressures. LV and RC volume initiations and terminations were all significantly higher in the COMF + 10 condition than in the COMF condition. These results are directly in line with the findings from Stathopoulos and Sapinenza (19) who used the same cue; however, they asked subjects to target a loudness 5 dB SPL higher than comfortable. Phonation onset time was shorter than in the COMF condition, indicating that subjects began speaking closer to the top of inhalation in the COMF + 10 condition. The mechanism of employing higher recoil pressures was used to the greatest extent in the COMF + 10 condition compared with the other two loud conditions, as demonstrated by the fact that the change from COMF for LV and RC volume initiations and terminations was greatest in the COMF + 10 condition (see Fig. 3). Furthermore, the COMF + 10 condition was the only loud condition in which LVIs and LVTs and RCVTs were significantly higher than in COMF. Use of higher LV and RC volume initiations and terminations was the predominant pattern for most of the trials in the COMF + 10 condition (see Fig. 6).

In the NOISE condition, RCVI and percent VC inspired were higher than in the COMF condition. These results suggest that speech was initiated at a higher LV, even though the LVI results were nonsignificant. Additionally, most trials demonstrated the pattern of increased LV and RC volume initiations and terminations during the NOISE condition (see Fig. 6). There was less change from COMF in LVIs and RCVIs in the NOISE condition than in the COMF + 10 condition, indicating that increased recoil pressure was used in the NOISE condition but not to the extent it was used in the COMF + 10 condition (see Fig. 3).

In the 2× COMF condition, LV and RC volume initiations and terminations did not change relative to the COMF condition, indicating that the lungs and RC were not more expanded when speech was initiated or terminated. Therefore, the primary mechanism for increasing loudness in the 2× COMF condition could not have been the use of higher recoil pressures and higher LVs. However, as in the COMF + 10 condition, phonation onset time was significantly shorter in the 2× COMF condition compared with the COMF condition, indicating that the subjects began to speak closer to the top of inhalation in the 2× COMF condition than in the COMF condition.

The larger reliance on the use of the higher recoil pressures in the COMF + 10 condition may have been a result of subject perception. Subjects may have perceived maintaining an SPL at nearly 90 dB as difficult and planned in advance to achieve this goal. Subjects may not have perceived the loudness target in the 2× COMF or NOISE conditions to be as high as the level in the COMF + 10 condition and, therefore, may not
have planned to need as much respiratory driving pressure even though they produced similar SPLs in all conditions. These data suggest that how the loudness target was perceived by the speaker affected the mechanisms used to support the loud speech and the neural control of the respiratory system.

Subjects did not utilize increased recoil pressures in the 2× COMF condition and utilized them to a lesser degree in the NOISE condition than in the COMF + 10 condition. However, they achieved the same overall increase in SPL. Therefore, another mechanism must have been used in the 2× COMF and NOISE conditions. In the 2× COMF condition, AbVI and AbVT were significantly lower than in the COMF condition, indicating that the Ab was more compressed during the 2× COMF condition. Because the Ab muscles are one of the major expiratory muscle groups (2, 3), the tucked position of the Ab suggests that the speakers generated higher expiratory muscle forces using their Ab muscles. The use of expiratory muscle tension to generate higher pressure for louder speech was most prevalent in the 2× COMF condition. The 2× COMF condition was the only loud condition in which there were significant changes from COMF in the Ab measurements (see Fig. 3). Furthermore, the pattern of decreased AbVI and AbVT was demonstrated for most trials in the 2× COMF condition, more than in the other two loud conditions (see Fig. 6).

In the NOISE condition, AbVI and AbVT were lower than in the COMF or COMF + 10 conditions with AbVI almost as low as in the 2× COMF condition, although the changes were nonsignificant (see Fig. 3). This suggests that subjects did use the mechanism of increasing expiratory muscle tension in the NOISE condition but not to the extent it was used in the 2× COMF condition. Furthermore, lower AbVI and AbVT were used on nearly as many trials in the NOISE condition as in the 2× COMF condition (see Fig. 6).

The biggest difference in engaging the abdomen between the 2× COMF and NOISE conditions is present in the AbVT data. The change in AbVT from COMF is much greater for 2× COMF than NOISE, whereas the change in AbVI from COMF is more similar between the two conditions (see Fig. 3). The large decrease in AbVT in the 2× COMF condition may suggest that subjects underestimated the amount of pressure that would be required to achieve the loudness target. They may have realized the need for more driving pressure as they moved through the utterance and used the Ab to generate higher expiratory muscle pressures. This is substantiated by the fact that the LVI and RCVI were lowest for the 2× COMF condition, indicating that subjects had less recoil pressure available in the 2× COMF condition than in the other two loud conditions (see Fig. 3).

It is interesting to note that there were no clear trends in the use of the Ab for the COMF + 10 condition. None of the Ab volume measurements (initiations, terminations, and excursions) changed significantly from COMF to COMF + 10, and there is no clear trend in the trial data for the Ab measurements (see Fig. 6). This finding is in line with previous studies that have demonstrated no contribution from the Ab toward increasing loudness (8, 19). However, given the findings for the 2× COMF and NOISE conditions in the present study, the conclusion that the Ab does not participate in increasing loudness appears to be related to how the increase in loudness is elicited.

In addition to the differences in how the increase in loudness was supported by the respiratory system for each condition, the NOISE condition stood out as different from the COMF + 10 and 2× COMF conditions in a number of ways. First, the NOISE condition was the only loud condition in which utterance duration was significantly longer, fewer syllables per second were produced, larger LVE were used, and a higher percent of VC was expended per syllable compared with COMF. Because the length of the utterances did not change across any of the conditions, all of these findings can be accounted for by a slower, more deliberate speech rate. Speakers may have perceived the need to use a slower speech rate in the NOISE condition to improve the intelligibility of the speech signal in the noise. These data suggest that goals for speech production were more complex in the NOISE condition than in the other two loud conditions, i.e., to improve intelligibility in addition to increasing loudness. This target for improved intelligibility in the NOISE condition was reflected in the respiratory kinematics.

Second, in the NOISE condition, the subjects combined the two proposed mechanisms for increasing loudness, higher recoil pressures, and more expiratory muscle tension. The use of a combined strategy may relate to the naturalness of the cue. The most natural cue was the NOISE condition since the subjects in the study had, presumably, spoken in noise previously in the course of their daily life. The 2× COMF condition may seem natural; however, speakers seldom think about doubling their loudness without additional environmental cues as to how much loudness increase is required (speaker-listener distance, increased room size, etc.). Therefore, the 2× COMF cue is not as natural as the NOISE cue. The combined respiratory strategy was demonstrated to the greatest extent in the NOISE condition.

Use of a combination of the two mechanisms may be advantageous for subjects since it may be less work for the system than using predominantly one mechanism. In using...
predominantly higher recoil pressures, a greater inspiratory effort must be expended to breathe to a higher LV and to control the high recoil pressures by checking the descent of the RC (3, 8). In using predominantly more expiratory muscle tension, a greater expiratory muscle effort must be expended to produce higher driving pressures for speech. A combination of these approaches would reduce both the inspiratory and expiratory muscle loads, spreading the work across a larger set of muscles. Furthermore, by not breathing to high LV or breathing farther below EEL, speakers stay closer to the mid-LV range, which has been suggested to be most efficient for speech production.

Last, in the NOISE condition, subjects inspired a greater percent of VC before each utterance than in the COMF and COMF + 10 conditions. This is particularly interesting since the COMF + 10 condition was the only condition in which the LV at which speech was initiated was higher. However, changes to RB have been reported as a result of noxious stimulation. The presence of noxious visual stimulation during RB has been shown to increase both tidal volume and frequency of breathing (14). The presence of saw-tooth noise during RB has been shown to increase ventilation in a group of individuals with high anxiety (15). Individuals may have breathed more deeply during the NOISE condition, resulting in a greater percent of VC inspired before an utterance. There are two possible explanations for the finding of greater VC inspired, but not greater LVI, in the NOISE condition. 1) Individuals may have let more air out before an utterance in the NOISE condition than in the COMF condition. 2) Alternately, individuals may have expired more after an utterance and before inhaling for the next utterance in the NOISE condition. This alternate hypothesis is supported by a study of decerebrate cats in which stimulation of the midbrain periaqueductal gray in noise resulted in louder vocalizations and increased laryngeal adductor and external oblique activation but no change in diaphragm activation (17).

The data from the present study suggest that speaking in noise elicits additional goals and different respiratory patterns than other conditions that were used to elicit an increase in SPL. However, the data do not support Winkworth and Davis’s (20) claims that speaking in noise does not elicit a patterned and consistent response from the respiratory system. There are clear patterns in the data, and the variability present does not match the level reported by Winkworth and Davis (see Fig. 6). The variability reported in their study may have been due to the reduction in all auditory feedback caused by delivering the noise through headphones. This methodology was altered in the present study since the noise was delivered via free field.

Although cerebral control of the respiratory system for speech production is not well understood, hypotheses can be made about the areas of the brain likely to be involved in the formation of internal targets and planning output to the respiratory muscles based on studies of the neural control of respiration. Cortical and subcortical activation, in addition to brain stem activation, would be expected since the respiratory system must be controlled voluntarily to achieve the goals of the speech task (16). Fink et al. (6) found volitional breathing to involve significantly higher activation of the supplementary motor area (SMA) than ventilator-controlled breathing. When resistance was added to increase the work of inspiration, there was also a significant increase in the activation of the primary motor area and the premotor area (6). These authors suggest that the increased activation with increased resistance might have been a result of “task-related changes in sensory input” (Ref. 6, p. 1304). The NOISE condition in the present study also involved a task-dependent change to sensory input and, therefore, generation of the resulting respiratory patterns may involve some of the same cortical areas suggested by Fink and colleagues (6). Further listening to continuous noise increases cerebral metabolism in the auditory areas and may increase the metabolic demand, leading to deeper breathing in noise (14).

McKay and colleagues (16) examined suprapontine activation during a voluntary hypernea breathing task. They found increased activation of the SMA bilaterally and the right premotor area (16). These authors suggested that the increase in activity in the SMA may relate to the learned nature of the task (16). This might explain the SMA activation in the Fink et al. (6) study as well since they asked subjects to breathe more deeply than normal. McKay et al. (16) suggest that the right premotor area activation may be due to an increased attentional requirement with the voluntary hypernea task. The SMA and premotor areas may be involved in the COMF + 10 condition since it is not a natural task and is likely to require a greater amount of attention than a more natural cue like the NOISE condition.

Also, in the COMF + 10 and NOISE conditions, subjects increased the volume at which speech was initiated, demonstrating preplanning of respiratory movements to achieve the internal target. This was not demonstrated in the 2× COMF condition, where subjects tended to continue to speak at lower LVs using predominantly expiratory muscle tension. The premotor area or SMA, which is known to be involved in motor planning, may have been more activated in the COMF + 10 and NOISE conditions more than in the 2× COMF condition.

The point of this study was not to suggest that one cue to increase loudness is better than another from a clinical perspective. However, an understanding of how cues affect respiratory function will assist in choosing the best cues for treatment. The results indicate that cues to increase loudness elicit different respiratory patterns. Because it is not possible to test every situation in which an individual may need to increase his/her loudness, it would be beneficial to treat individuals using multiple cues to ensure several patterns of respiratory function are supported by therapy. Furthermore, it is important to realize that results from studies of respiratory kinematics cannot be compared across studies without considering the cue used to elicit the increase in loudness. Last, it is important to consider the efficiency of the patterns elicited by these cues, from a work perspective, when planning a treatment. The kinematic patterns elicited in the NOISE condition appeared to be the most efficient and required the least muscular effort from the speaker. The kinematic pattern elicited by the COMF + 10 condition was also efficient in that recoil pressures were used to a great extent, reducing the expiratory muscle effort. The kinematic patterns elicited by the 2× COMF condition appeared to be the least efficient of the cues used in the present study since the large majority of respiratory driving pressure was generated by increasing expiratory muscle tension. Speech in the 2× COMF condition was produced...
at LVs where recoil pressures are lower than in the COMF + 10 and NOISE conditions, resulting in a greater amount of pressure to be generated by muscle tension.

In summary, the data from the present study suggest that different cues to increase loudness result in different internal targets and that the neural control of the respiratory system for speech is sensitive to changes in the speaker’s internal target. For example, in the most natural the elicitation method (NOISE), the subjects used a combined respiratory approach: both increased recoil pressures and increased expiratory muscle tension. Furthermore, changes to speech rate that accompanied changes in SPL in the NOISE condition were reflected in the respiratory kinematics. The control of the respiratory system also seemed to reflect speakers’ perceptions or expectations. One example of this was in the COMF system also seemed to reflect speakers’ perceptions or expectations. The control of the respiratory muscles in breathing. Furthermore, changes to speech rate that accompanied changes in SPL in the NOISE condition were reflected in the respiratory kinematics. The control of the respiratory system also seemed to reflect speakers’ perceptions or expectations. One example of this was in the COMF + 10 condition where a strategy of using primarily higher recoil pressures was demonstrated possibly because of subjects’ expectations of difficulty in reaching this loudness target. Another example of this was in the 2× COMF condition where a primarily Ab strategy was used, possibly because subjects misjudged the amount of respiratory drive needed for this condition. In future studies, it would be beneficial to study the effects of these cues across a continuum. For example, does a lower level of noise induce the same changes to respiratory patterns and speech rate?

GRANTS

This research was funded by National Institute on Deafness and Other Communication Disorders Grant 1R03 DC-05731-01.

REFERENCES