Published in final edited form as:

J Appl Physiol. 2005 June ; 98(6): 2177–2184. doi:10.1152/japplphysiol.01239.2004.

Changes to Respiratory Mechanisms during Speech as a Result of Different Cues to Increase Loudness

Jessica E. Huber, Bharath Chandrasekaran, and John J. Wolstencroft

Purdue University, Department of Speech, Language, and Hearing Sciences, West Lafayette, IN 47907

Abstract

The purpose of the present study was to determine if 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 thirty 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 (SPL), 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 SPL. 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 SPL, 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 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.

Keywords

Speaking in noise; Respiratory kinematics; Motor planning in speech

INTRODUCTION

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 to take advantage of higher recoil pressures (4,8,19), although all speakers do not increase their lung volume to the same degree (19). A second mechanism, more often used by speakers closer to or below end expiratory level, is to increase expiratory muscle tension in order 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 (1997) 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 which reported a pattern of initiating speech at higher lung volumes. 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 (1997) 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, a) asking subjects to speak twice or four times as loud as comfortable (4) and b) 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 lung volume initiations (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 (1997) suggest this may not be true. Their data demonstrate 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 which *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 in order 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 current study, the goal is to be louder. The internal target results in a scaling up of the movement plans which 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 in order 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 (2000), 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 percept of the object, and the internal target was reflected in the movement parameters for reaching and grasping. Since 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 if 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 current study is also important from a clinical perspective in speech therapy. There are a number of speech disorders which result in reduced loudness levels, such as Parkinson's disease. Individuals with symptoms of low vocal loudness often undergo therapies which cue them, in many ways, to increase their loudness. Since 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.

MATERIAL AND METHODS

Subjects

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. 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 non-smoking for the past five years. They had a body mass index (BMI) 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 HL for octave frequencies between 250 and 8000 Hz, bilaterally, completed in a quiet room. Subjects had normal vital capacity (VC), forced vital capacity (FVC), and forced expiratory volume in one second (FEV1.0) defined as equal to or better than 80% of expected values based on age, gender, height, weight, and ethnicity. Lung capacities were measured using a digital spirometer (VacuMed Discovery Handheld Spirometer).

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 inches away. The conditions were as follows:

1. COMF: Subjects were instructed to "read the sentence at your comfortable loudness and pitch."
2. 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 which faced the subject. The feedback was provided continuously to the subject.

3. 2XCOMF: 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.
4. NOISE: Multi-talker noise (AUDiTEC of St. Louis) 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 inches in front of the subjects.

For each condition, the shorter sentence was completed first, and each sentence was said fifteen times consecutively. The COMF condition was always completed first. The order of the three loud conditions (COMF+10, 2XCOMF, 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, and abdomen. 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. The VC measured as a part of the inclusion criteria was used and lung volume measures were expressed as a percentage of the largest VC produced. Vital capacity maneuvers were also completed at least 3 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 rib cage capacity (RCC) since the rib cage moves maximally during a VC maneuver. To determine abdominal 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 end expiratory level (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 the total ABC. At least three steady cycles of rest breathing were collected prior to the start of each trial of the maximum capacity maneuvers.

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 subject'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 (1). 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.

Respiratory kinematic data were transduced via respiratory inductive plethysmography using the RespiTrace system (Ambulatory Monitoring, Inc.). An elastic band was placed around the rib cage, just under the axilla to transduce movements of the rib cage. 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

2000 Hz. Data from a second microphone was 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 ten consecutive sentences which were produced without error were chosen for analysis. Sound Pressure Level (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.

Since lung volume change reflects the combined effect of changes in the rib cage (RC) and abdominal (AB) volumes (11), the sum of the RC and AB signals was computed and corrected for the respective RC and AB contributions to lung volume (LV) change. Two non-speech tasks were used to determine the RC and AB contributions to LV change. Subjects were instructed to relax and data was collected for two 45-second periods of rest breathing (RB). Data was also collected for three 45-second 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 rest breathing were collected prior to the start of the each SLB data collection period. Lung volume data was collected during the RB and SLB tasks using a spirometer (VacuMed Universal Ventilation Meter). This spirometer has a very small dead space. These data were digitized along with the respiratory kinematic data at 2000 Hz.

The data from the spirometer (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 against the original SP signal for a “speech-like breathing” trial. The estimated lung volume signal was then computed for each point during the sentence production tasks using the formula Lung Volume (LV) = $k_1(RC) + k_2(AB)$.

LV, RC, and AB initiations and terminations were measured relative to end expiratory level (EEL). EEL was measured from troughs of three steady rest breaths prior to 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 Figure 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. 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 Figure 1, line A – line B).

Percent VC expended per syllable was measured by dividing the lung volume excursion by the number of syllables for each utterance. Percent VC inspired was measured from the lung volume 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 Figure 1, point D – 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 Figure 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 analyses of variance (ANOVA). The within factor was loudness (condition) and the between factor was sex. Table 1 provides a summary of the statistical results. Tukey HSD 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 $p < 0.01$.

To test for a learning effect across the ten 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 $p < 0.01$, as for the ANOVAs. There were no significant differences between the two means for any of the measurements, suggesting there was no significant learning effect.

Inter-measurer reliability was completed on 2 male and 2 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 inter-measurer 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 (standard error (SE) = 0.55 dB) for COMF, 89 dB (SE = 0.61 dB) for COMF+10, 88 dB (SE = 0.73 dB) for 2XCOMF, 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 Figure 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 as compared to the COMF condition (see Figure 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 2XCOMF conditions as compared to the COMF condition (see Figure 2). Respiratory Kinematic Measurements:

For lung volume initiation (LVI), there were significant condition and sex effects, but no interaction effect. LVI was significantly higher in the COMF+10 condition as compared to the COMF condition (see Figure 3). LVI was significantly higher for the women as compared to the men.

For lung volume termination (LVT), there were significant condition and sex effects, but no interaction effect. LVT was significantly higher for the COMF+10 condition as compared to

the COMF condition (see Figure 3). LVT was significantly higher for the women than for the men.

For lung volume 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 Figure 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 Figure 5).

For percent VC expended per syllable, there was a significant condition effect, but no sex or interaction effects. Percent VC expended per syllable was significantly higher in the NOISE condition as compared to the COMF condition (see Figure 5).

For rib cage 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 as compared to COMF condition (see Figure 3).

For rib cage volume termination (RCVT), there was a significant condition effect, but no sex or interaction effects. RCVT was significantly higher in the COMF+10 condition as compared to the COMF condition (see Figure 3).

For rib cage volume excursion (RCVE), there was a significant condition effect, but no sex or interaction effects. RCVE was significantly larger in all loud conditions as compared to the COMF condition (see Figure 4).

For abdominal volume initiation (ABVI), there was a significant condition effect, but no sex or interaction effects. ABVI was significantly lower in the 2XCOMF condition as compared to the COMF condition (see Figure 3).

For abdominal volume termination (ABVT), there was a significant condition effect, but no sex or interaction effects. ABVT was significantly lower in the 2XCOMF condition, as compared to the COMF condition (see Figure 3).

For abdominal volume excursion (ABVE), there were no condition, sex, or interaction effects (see Figure 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, about 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 lung volume in order to utilize higher recoil pressures. Lung and rib cage 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 Sapienza (1997) 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 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 to the other two loud conditions, as demonstrated by the fact that the change from COMF for lung and rib cage volume initiations and terminations was greatest in the COMF+10 condition (see Figure 3). Further, the COMF+10 condition was the only loud condition in which lung volume initiations and terminations and rib cage terminations were significantly higher than in COMF. Use of higher lung and rib cage volume initiations and terminations was the predominant pattern for most of the trials in the COMF+10 condition (see Figure 6).

In the NOISE condition, rib cage volume initiation and percent vital capacity inspired were higher than in the COMF condition. These results suggest that speech was initiated at a higher lung volume, even though the lung volume initiation results were non-significant. Additionally, most trials demonstrated the pattern of increased lung and rib cage volume initiations and terminations during the NOISE condition (see Figure 6). There was less change from COMF in lung and rib cage volume initiations in the NOISE condition than in the COMF+10 condition, indicating increased recoil pressure was used in the NOISE condition, but not to the extent it was used in the COMF+10 condition (see Figure 3).

In the 2XCOMF condition, lung and rib cage volume initiations and terminations did not change relative to the COMF condition, indicating that the lungs and rib cage were not more expanded when speech was initiated or terminated. Therefore, the primary mechanism for increasing loudness in the 2XCOMF condition could not have been the use of higher recoil pressures and higher lung volumes. However, as in the COMF+10 condition, phonation onset time was significantly shorter in the 2XCOMF condition, as compared to the COMF condition, indicating the subjects began to speak closer to the top of inhalation in the 2XCOMF 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 2XCOMF 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 as 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 2XCOMF 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 2XCOMF and NOISE conditions. In the 2XCOMF condition, abdominal volume initiations and terminations were significantly lower than in the COMF condition, indicating that the abdomen was more compressed during the TAL condition. Since the abdominal muscles are one of the major expiratory muscle groups (2,3), the tucked position of the abdomen suggests that the speakers generated higher expiratory muscle forces using their abdominal muscles. The use of expiratory muscle tension to generate higher pressure for louder speech was most prevalent in the 2XCOMF condition. The 2XCOMF condition was the only loud condition in which there were significant changes from COMF in the abdominal measurements (see Figure 3). Further, the pattern of decreased abdominal volume initiations and terminations was demonstrated for most trials in the 2XCOMF condition, more than in the other two loud conditions (see Figure 6).

In the NOISE condition, abdominal volume initiations and terminations were lower than in the COMF or COMF+10 conditions, with abdominal volume initiations almost as low as in the 2XCOMF condition, although the changes were non-significant (see Figure 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 2XCOMF condition. Further, lower abdominal volume initiations and terminations were used on nearly as many trials in the NOISE condition as in the 2XCOMF condition (see Figure 6).

The biggest difference in engaging the abdomen between the 2XCOMF and NOISE conditions is present in the abdominal volume termination data. The change in abdominal volume terminations from COMF is much greater for 2XCOMF than NOISE, whereas the change in abdominal volume initiations from COMF is more similar between the two conditions (see Figure 3). The large decrease in abdominal volume terminations in the 2XCOMF condition may suggest that subjects underestimated the amount of pressure which 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 abdomen to generate higher expiratory muscle pressures. This is substantiated by the fact that the lung and rib cage volume initiations were lowest for the 2XCOMF condition, indicating that subjects had less recoil pressure available in the 2XCOMF condition than in the other two loud conditions (see Figure 3).

It is interesting to note that there were no clear trends in the use of the abdomen for the COMF +10 condition. None of the abdominal 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 abdominal measurements (see Figure 6). This finding is in line with previous studies which have demonstrated no contribution from the abdomen toward increasing loudness (8,19). However, given the findings for the 2XCOMF and NOISE conditions in the present study, the conclusion that the abdomen 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 2XCOMF 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 lung volume excursions were used, and a higher percent of vital capacity was expended per syllable as compared to COMF. Since 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 in order 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, 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 2XCOMF 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 2XCOMF 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 lung volume and to control the high recoil pressures by checking the descent of the rib cage (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. Further, by not breathing to high lung volumes or breathing farther below EEL, speakers stay closer to the mid-lung volume range which has been suggested to be most efficient for speech production.

Last, in the NOISE condition, subjects inspired a greater percent of vital capacity 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 lung volume at which speech was initiated was higher. However, changes to rest breathing have been reported as a result of noxious stimulation. The presence of noxious visual stimulation during rest breathing has been shown to increase both tidal volume and frequency of breathing (14). The presence of saw-tooth noise during rest breathing 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 vital capacity inspired prior to an utterance. There are two possible explanations for the finding of greater vital capacity inspired, but not greater lung volume initiation, in the NOISE condition. Individuals may have let more air out prior to an utterance in the NOISE condition than in the COMF+10 condition. This possibility is supported by the fact that phonation onset time was not significantly shorter than COMF in the NOISE condition, as it was in the COMF+10 condition. Alternately, individuals may have expired more after an utterance, 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 current study suggest that speaking in noise elicits additional goals and different respiratory patterns than other conditions which were used to elicit an increase in SPL. However, the data do not support Winkworth and Davis's (1997) 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 Figure 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 current study since the noise was delivered via free-field.

While 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 brainstem activation, would be expected since the respiratory system must be controlled voluntarily in order to achieve the goals of the speech task (16). Fink et al. (1996) 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" (6, p. 1304). The NOISE condition in the current 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 (1996). 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 (2003) 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 (1996) study as well since they asked subjects to breathe more deeply than normal. McKay et al. (2003) 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 pre-planning of respiratory movements to achieve the internal target. This was not demonstrated in the 2XCOMF condition where subjects tended to continue to speak at lower lung volumes using predominantly expiratory muscle tension. The premotor area or SMA, which are known to be involved in motor planning, may have been more activated in the COMF+10 and NOISE conditions more than in the 2XCOMF 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 cues to increase loudness elicit different respiratory patterns. Since 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. Further, it is important to realize that results from studies of respiratory kinematics can not 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 2XCOMF condition appeared to be 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. Also, speech in the 2XCOMF condition was produced at lung volumes 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 current 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. Further, changes to speech rate which 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 2XCOMF condition where a primarily abdominal 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?

Acknowledgements

Grants:

This research was funded by the National Institutes of Health, National Institute on Deafness and Other Communication Disorders, grant # 1R03DC05731-01.

References

1. Boersma, P.; Weenink, D. Praat. Vol. 4.1. Amsterdam: Institute of Phonetic Sciences; 2003.
2. Campbell EJM. An electromyographic study of the role of the abdominal muscles in breathing. *J Physiol* 1952;117:223–233.
3. Draper MH, Ladefoged P, Whitteridge D. Respiratory muscle in speech. *J Speech and Hear Res* 1959;2:16–27. [PubMed: 13655288]
4. Dromey C, Ramig LO. Intentional changes in sound pressure level and rate: Their impact on measures of respiration, phonation, and articulation. *J Speech Lang Hear Res* 1998;41:1003–1018. [PubMed: 9771624]
5. Treatment of Overweight and Obesity in Adults. Expert Panel on the Identification E. Executive summary of the clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. *Archives of Internal Medicine* 1998;158:1855–1867. [PubMed: 9759681]
6. Fink GR, Corfield DR, Murphy K, Kobayashi I, Dettmers C, Adams L, Frackowiak RSJ, Guz A. Human cerebral activity with increasing inspiratory force: a study using positron emission tomography. *J Appl Physiol* 1996;81:1295–1305. [PubMed: 8889766]
7. Gentilucci M, Benuzzi F, Bertolani L, Daprati E, Gangitano M. Language and motor control. *Exp Brain Res* 2000;133:468–490. [PubMed: 10985682]
8. Hixon TJ, Goldman MD, Mead J. Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. *J Speech and Hear Res* 1973;16:78–115. [PubMed: 4267384]
9. Hoit JD, Hixon TJ. Age and speech breathing. *J Speech and Hear Res* 1987;30:351–366. [PubMed: 3669642]
10. Huber JE, Stathopoulos ET. Respiratory and laryngeal responses to an oral air pressure bleed during speech. *J Speech Lang Hear Res* 2003;46:1207–1220. [PubMed: 14575353]
11. Konno K, Mead J. Measurement of the separate volume changes of the rib cage and the abdomen during breathing. *J Appl Physiol* 1967;22:407–422. [PubMed: 4225383]
12. Lane H, Tranel B. The Lombard Sign and the role of hearing in speech. *J Speech and Hear Res* 1971;14:677–709.
13. Lieberman, P. Intonation, Perception, and Language. Cambridge, MA: The M.I.T. Press; 1967. Physiologic, acoustic, and perceptual criteria.
14. Mador MJ, Tobin MJ. Effect of alterations in mental activity on the breathing pattern of healthy subjects. *Am Rev Respir Dis* 1991;144:481–487. [PubMed: 1892283]
15. Masaoka Y, Homma I. Expiratory time determined by individual anxiety levels in humans. *J Appl Physiol* 1999;86:1329–1336. [PubMed: 10194219]
16. McKay LC, Evans KC, Frackowiak RSJ, Corfield DR. Neural correlates of voluntary breathing in humans. *J Appl Physiol* 2003;95:1170–1178. [PubMed: 12754178]
17. Nonaka S, Takahashi R, Enomoto K, Katada A, Unno T. Lombard reflex during PAG-induced vocalization in decerebrate cats. *Neuroscience Research* 1997;29:283–289. [PubMed: 9527619]
18. Pick HL Jr, Siegel GM, Fox PW, Garber SR, Kearney JK. Inhibiting the Lombard effect. *J Acoust Soc Am* 1989;85:894–900. [PubMed: 2926004]
19. Stathopoulos ET, Sapienza CM. Developmental changes in laryngeal and respiratory function with variations in sound pressure level. *J Speech Lang Hear Res* 1997;40:595–614. [PubMed: 9210117]

20. Winkworth AL, Davis PJ. Speech breathing and the Lombard effect. *J Speech Lang Hear Res* 1997;40:159–169. [PubMed: 9113867]

Figure 1.

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

Figure 2.

Timing measurements: mean difference from comfortable (COMF) for the three loud conditions (COMF+10, 2XCOMF, and NOISE). Lines show standard errors; asterisks indicate significant changes from COMF.

Figure 3.

Initiation and termination measurements: mean difference from comfortable (COMF) for the three loud conditions (COMF+10, 2XCOMF, and NOISE). Lines show standard errors; asterisks indicate significant changes from COMF. 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.

Figure 4.

Excursion measurements: mean difference from comfortable (COMF) for the three loud conditions (COMF+10, 2XCOMF, and NOISE). Lines show standard errors; asterisks indicate significant changes from COMF. LVE = lung volume excursions; RCVE = rib cage volume excursion, ABVE = abdominal volume excursion.

Figure 5.

Percent vital capacity expended and percent vital capacity inspired measurements: mean difference from comfortable (COMF) for the three loud conditions (COMF+10, 2XCOMF, and NOISE). Lines show standard errors; asterisks indicate significant changes from COMF.

Figure 6.

Histogram of respiratory kinematic initiations and terminations showing the number of trials in which the COMF condition was higher or lower the COMF+10, 2XCOMF, and NOISE conditions, respectively.

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

Measures	Condition (3, 84)		Sex (1, 28)		Condition X Sex (3, 84)	
	F	p	F	p	F	p
Sound Pressure Level	122.51	0.000*	0.27	0.606	0.52	0.670
Duration	4.57	0.005*	2.76	0.108	0.76	0.522
Syllables Per Second	4.63	0.005*	3.81	0.061	1.16	0.329
Phonation Onset Time	5.34	0.002*	0.22	0.645	2.42	0.072
Lung Volume Initiation	6.65	0.000*	17.10	0.000*	2.46	0.068
Lung Volume Termination	5.00	0.003*	14.73	0.001*	2.28	0.085
Lung Volume Excursion	7.87	0.000*	3.82	0.061	0.60	0.617
Percent Vital Capacity Inspired	9.30	0.000*	2.60	0.118	3.86	0.012
Percent Vital Capacity	8.89	0.000*	3.23	0.083	1.11	0.351
Expended Per Syllable						
Rib Cage Volume Initiation	8.17	0.000*	5.31	0.029	2.51	0.064
Rib Cage Volume Termination	4.25	0.008*	5.96	0.021	2.82	0.044
Rib Cage Volume Excursion	12.51	0.000*	0.14	0.706	0.59	0.622
Abdominal Volume Initiation	6.44	0.001*	0.20	0.657	1.31	0.28
Abdominal Volume Termination	5.29	0.002*	0.06	0.804	1.60	0.196
Abdominal Volume Excursion	0.44	0.723	1.88	0.181	1.19	0.319

Degrees of freedom in parentheses. F = F-ratio, p = level of significance.

Asterisk indicates significance at the $p < 0.01$ level.