# REPORT DOCUMENTATION PAGE

*Form Approved OMB No. 0704-0188*



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | November, 1990 | |

**4. TITLE AND SUBTITLE**
A NOTE ON LOUD AND LOMBARD SPEECH

**5. FUNDING NUMBERS**
WU - 2313V301

**6. AUTHOR(S)**
Z. S. BOND AND THOMAS J. MOORE

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Department of Linguistics
Ohio University
Athens OH

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Armstrong Aerospace Medical Research Laboratory
Wright-Patterson AFB OH 45433-6573

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
AMRL-TR-90-095

**11. SUPPLEMENTARY NOTES**
AAMRL/Contact: Dr. Thomas J. Moore, Det 1, AL/BBA
Tel: (513) 255-3607

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution is unlimited.

**12b. DISTRIBUTION CODE**
A

**13. ABSTRACT** *(Maximum 200 words)*
The purpose of this report is to compare speech which is loud as a consequence of noise exposure (Lombard speech) with speech which is loud deliberately. One male speaker was recorded in six speaking conditions: ambient noise, high noise, and intentionally loud speech, all three recorded with a boom microphone and while wearing an oxygen mask. Lombard speech and deliberately loud speech shared more similarities than differences and appear to result from the same speech production mechanisms. Both were produced with more effort so that energy and fundamental frequency increased. Both were produced with a wider mouth opening so that formants, particularly F1, shifted. The oxygen mask minimized changes.



**14. SUBJECT TERMS**
Noise Effects
Speech
Voice Communications

**15. NUMBER OF PAGES**
4

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | UL |

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std Z39-18
298-102

# A NOTE ON LOUD AND LOMBARD SPEECH

Z. S. Bond and Thomas J. Moore

Department of Linguistics, Ohio University, Athens, Ohio
Armstrong Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio

ABSTRACT

The purpose of this report is to compare speech which is loud as a consequence of noise exposure (Lombard speech) with speech which is loud deliberately. One male speaker was recorded in six speaking conditions: ambient noise, high noise, and intentionally loud speech, all three recorded with a boom microphone and while wearing an oxygen mask. Lombard speech and deliberately loud speech shared more similarities than differences and appear to result from the same speech production mechanisms. Both were produced with more effort so that energy and fundamental frequency increased. Both were produced with a wider mouth opening so that formants, particularly F1, shifted. The oxygen mask minimized changes.

## I. INTRODUCTION

In the past few years, it has become clear that speech produced in an environment in which the speaker hears high levels of noise differs appreciably from speech produced in a benign environment. Not only does the pitch and loudness (that is fundamental frequency and energy) increase, but there is an increase in the proportion of high frequency energy in the speech spectrum and the vowel space as defined by F1 and F2 changes; particularly noticeable is a shift upwards of F1. There may be a downward shift of F3 [1,2,4].

In this past work describing changes in speech as a consequence of noise exposure, or Lombard speech, the speakers have typically been placed in a noisy environment and asked to read a list of words, phrases or sentences. Any changes in their speech have been assumed to be a by-product of noise exposure. They have typically not been asked to speak loudly or clearly.

The purpose of this report is to compare speech which is loud as a consequence of noise exposure with speech which is loud deliberately, as a result of speaking style. The question of interest is: Is loud speech the same, whether deliberately produced or a by-product of other speaking circumstances.

## II. METHOD

The speaker was a young male who had participated in the study reported by Bond, Moore, and Gable [1]. He was asked to return and to record a list of spondee words in an ambient condition and a 95 dB pink noise exposure condition, as described previously. The speaker was also asked to record the same materials while speaking deliberately loudly. He could see a VU meter showing his speech level; in addition, he was asked to imagine his audience at a distance. The speaker was recorded with a boom microphone and while wearing a standard AF flight helmet and oxygen mask equipped with a noise canceling microphone (M-101). There were six speaking conditions of interest: ambient noise, high noise, and intentionally loud speech, all three recorded while wearing a boom microphone and also while wearing an oxygen mask.

Data analysis was conducted using SPIRE on the Symbolics 3670 computer, as described in the previous publication. Measurements included word and syllable nucleus durations, fundamental frequency, and the frequencies and amplitudes of the first three formants.

## III. RESULTS

### 3.1 Duration

The average durations of all words as spoken in the six speaking conditions are given in Table I. In this and following tables and figures, the speaking conditions are identified by the following abbreviations: A=ambient, L=deliberately loud, N=noise, MA=wearing oxygen mask, ML=loud with mask, MN=noise and mask.

The variation in word durations was relatively small. In the ambient condition, the average word duration was

731 msc. Words were longest in the noise condition, 831 msc., and shortest in the mask/loud condition, 680 msc. Since the standard deviations ranged from 91 to 112 msc., these duration differences were within one standard deviation of each other.

Table I. Word durations (mean and standard deviation) in msc.

| CONDITION | MEAN | SD |
|---|---|---|
| A | 731 | 100 |
| L | 735 | 112 |
| N | 831 | 108 |
| MA | 773 | 93 |
| ML | 680 | 91 |
| MN | 750 | 100 |

The durations of syllable nuclei are given in Table II. The second syllable of the spondee words was invariably longer than the first syllable, probably because the second syllable was in utterance-final position and subject to pre-boundary lengthening. In the ambient condition, the average duration of the vocalic portion of the first syllable was 156 msc., and of the second syllable 234 msc., the shortest durations. The standard deviations were within similar ranges for each of the syllables. As in the case of words, the durational differences between the speaking conditions increased by less than one standard deviation from the ambient condition.

Table II. Durations of the syllable nuclei in msc.

| CONDIT'N | S1:MEAN | SD | S2:MEAN | SD |
|---|---|---|---|---|
| A | 156 | 33 | 234 | 65 |
| L | 178 | 43 | 257 | 60 |
| N | 172 | 35 | 275 | 60 |
| MA | 170 | 42 | 248 | 57 |
| ML | 168 | 42 | 235 | 58 |
| MN | 172 | 40 | 248 | 63 |

Shulman [3] reported small durational differences between speech produced loudly and speech produced at normal levels. In spite of the fact that loud speech seems to require larger oral cavity openings and hence greater displacement, speakers apparently used greater velocity of movement to compensate for the displacement. Shulman based his conclusions on studies of articulatory dynamics. His findings are supported by the acoustic measurements reported here. Assuming that the speaker was also using greater displacement of the articulators in producing loud speech, he must have been adjusting the rate of movement so that durational differences were negligible.

In the mask conditions, word durations were more similar to each other than the durations produced in the no-mask conditions, perhaps resulting from movement restrictions which wearing the oxygen mask imposes.

3.2 Fundamental Frequency

An increase in fundamental frequency almost always accompanies Lombard speech. As can be seen from Table III, fundamental frequency increased when the speaker was speaking in noise, by 33% in the first syllable and by 20% in the second syllable. Fundamental frequency also increased in loud speech, though by somewhat smaller percentages, 20% in the first syllable and 19% in the second syllable. When the speaker was wearing the oxygen mask, the fundamental frequency changes were smaller for both Lombard speech and loud speech.

Table III. F0 at the mid-point of the two syllables (Hz).

| CONDIT'N | S1:MEAN | SD | S2:MEAN | SD |
|---|---|---|---|---|
| A | 131 | 6 | 104 | 3 |
| L | 158 | 9 | 123 | 5 |
| N | 175 | 8 | 126 | 5 |
| MA | 129 | 7 | 109 | 3 |
| ML | 150 | 7 | 119 | 4 |
| MN | 135 | 8 | 114 | 4 |

3.3 Total Energy

The energy of speech has also been found to increase consistently when the speaker is in a noisy environment. Both loud and Lombard speech showed an increase, 2.2 dB in the first syllable, more in the second. With the mask, the differences in energy were considerably reduced. In the mask condition, loud and Lombard speech differed from ambient speech by less than 1 dB in the first syllable. The second syllable in the mask / loud condition decreased in total energy. There was a tendency in loud speech for the first syllable to show considerably more change than the second syllable, as if the speaker were putting most of his energy into the stressed syllable. These data are given in Table IV.

Table IV. Changes in total energy in dB from ambient condition to loud and Lombard speech conditions.

| CONDIT'N | in S1 | in S2 |
|---|---|---|
| L | 2.2 | 2.8 |
| N | 2.2 | 6.1 |
| ML | 0.4 | -1.7 |
| MN | 0.9 | 2.8 |

### 3.4 Formant Frequencies

The vowel space as defined by the first two formants is given in Fig. 1. The data represent the four vowels found at the extremes of the vowel quadrilateral /i, ae, a, u/. In the noise condition, the vowel space appeared to constrict in comparison with the formant values found in the ambient condition, particularly in the F1 plane. Averaging over all tokens of the four vowels, the first formant in Lombard speech was almost 80 Hz higher than in the ambient condition. The tendency seemed to be general, even though high vowels were affected more than low vowels. For high vowels, F1 increased by 105 Hz while the increase was 54 Hz for low vowels. Changes in F2 suggested a somewhat fronted tongue position, again most marked for the two high vowels.

The vowel formants of loud speech exhibited similar effects for the two high vowels /i, u/. For the two low vowels, the first formant in loud speech decreased somewhat, having values very similar to the values found in the oxygen mask conditions. Averaged over all tokens of the two high vowels, F1 increased by 109 Hz in loud speech in comparison with speech in the ambient condition. For the low vowels, F1 decreased by 46 Hz. F1 differences were reduced in the mask conditions, perhaps because F1 was relatively high in the mask condition.

Shulman [3] reported that previous acoustical measurements of vowel formants in loud speech showed a substantial increase in F1 which may be a consequence of a lowered jaw. These data support his interpretation, particularly for the high vowels.

The third formant, averaged over all tokens of all vowels, was almost 90 Hz lower in loud speech than in speech produced in the ambient condition. It was also somewhat lowered in all mask conditions in comparison with the ambient condition. However, there was little effect of noise on F3 for this speaker.

## IV. CONCLUSION

The general impression of the data was that Lombard speech which is inadvertently loud and deliberately loud speech resulted from the same speech production mechanisms. Both kinds of speech were produced with more effort so that energy and fundamental frequency increased. Both were produced with a wider mouth opening--a lower jaw position--so that formants, particularly F1, shifted. Wearing the oxygen mask minimized the fundamental frequency and energy changes.

There were only two differences between loud and Lombard speech. In loud speech, the fundamental frequency and energy of the first syllable was affected more than the second syllable, suggesting that the speaker concentrated vocal effort on the stressed first syllable. The frequency of the first formants of low vowels did not increase.

Since the data represent the speech of one speaker, these differences must be interpreted with caution. They may be characteristic of this speaker rather than general tendencies. Furthermore, it is possible that the speaker did not produce loud speech in the same way each time. He was simply asked to be loud and may have varied his interpretation of loudness or even forgotten to be loud on occasion. To investigate the details of loud speech with more certainty, it would be necessary to devise an experimental protocol in which the speaker has intrinsic motivation to speak loudly, perhaps because a listener is placed at some distance from the speaker. Until loud speech is elicited in a more realistic setting, comparisons between Lombard and loud speech have to be tentative. Even so, the two speaking conditions have many more similarities than differences.

## REFERENCES

[1] Z. S. Bond, T. J. Moore, and B. Gable, Acoustic-phonetic characteristics of speech produced in noise and while wearing an oxygen mask, Journal of the Acoustical Society of America, 85, 1989, 907-912.

[2] Z. S. Bond, T. J. Moore, and K. McCreight, Acoustic characteristics of sentences produced in noise, AAMRL-TR-89-036, 1989,Armstrong Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio.

[3] R. Shulman, Articulatory dynamics of loud and normal speech, Journal of the Acoustical Society of America, 85, 1989, 295-312.

[4] W. V. Summers, D. B. Pisoni, R. H. Bernacki, R. I. Pedlow, and M. A. Stokes, Effects of noise on speech production: Acoustic and perceptual analyses, Journal of the Acoustical Society of America, 84, 1988, 917-928.

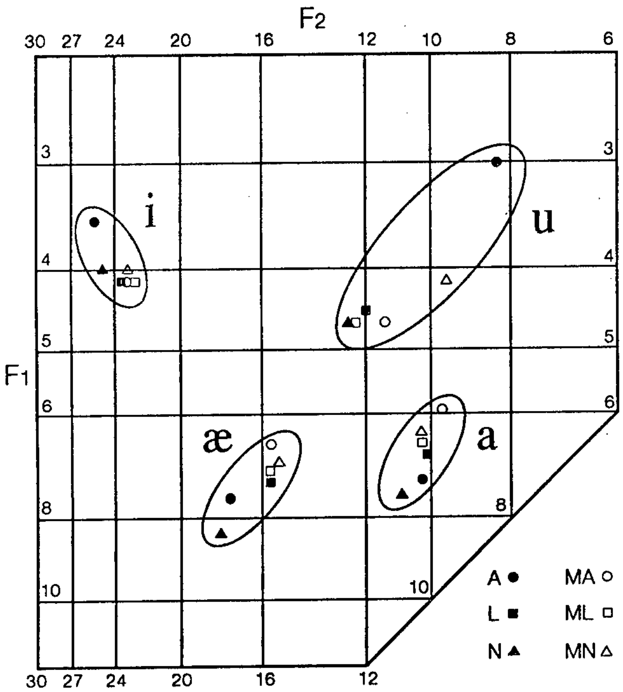


Fig. 1. Vowel space as defined by the frequencies of the first formant (vertical axis) and the second formant (horizontal axis). Frequency is given in hundreds of Hz.