Long-term Average Speech Spectrum at the Chest-level Microphone Location

Enregistrements au niveau de la poitrine du spectre moyen de la parole (SMP)

Leonard E. Cornelisse, Jean-Pierre Gagné, and Richard C. Seewald
Hearing Health Care Research Unit
Department of Communicative Disorders
The University Of Western Ontario

Keywords: Speech spectrum, FM systems, hearing aids, long-term average spectrum of speech, chest-level microphone placement

Abstract

The long-term average spectrum of speech (LTASS) obtained from 10 adult females was measured at two different recording positions: 30 cm directly in front of the talker (reference position); and 20 cm below the talker's mouth, at the position of an FM transmitter microphone (chest-level position). Results indicated significant differences in the overall level and frequency spectra between the LTASS obtained at each microphone location. A transfer function for the chest-level microphone position was calculated from the data of the present investigation. A representation of the LTASS that approximates the spectrum received at the chest-level microphone of an FM transmitter was derived by adding the transfer function for a chest-level microphone and a correction for vocal effort to the LTASS proposed for hearing aids. The derived spectrum closely approximates a representation of the LTASS suggested for application in the fitting of FM systems.

Résumé

Le spectre moyen de la parole (SMP) obtenu de 20 femmes adultes a été mesuré à deux positions d'enregistrement : à 30 centimètres directement en avant de la personne qui parle (position de référence) et à 20 centimètres sous la bouche de la personne qui parle (au niveau de la poitrine). Les résultats ont indiqué des différences significatives dans l'ensemble du spectre de niveau et du spectre de fréquence entre le SMP obtenu à chacune des positions du microphone. Une fonction de transfert de la position du microphone au niveau de la poitrine a été calculée en fonction des données de l'enquête. Une représentation du SMP qui équivaut approximativement au spectre reçu au microphone placé au niveau de la poitrine d'un transmetteur MF a été obtenu en ajoutant la fonction de transfert pour un microphone au niveau de la poitrine et une correction de l'effort vocal au SMP proposé pour les prothèses auditives. Le spectre obtenu est presque équivalent au SMP que l'on propose d'appliquer dans l'ajustement des systèmes.

The long-term average spectrum of speech (LTASS) is a direct measure of the time-averaged sound pressure level of speech as a function of frequency. Group mean representations of the LTASS have been used in the calculation of required gain in hearing aid prescription procedures (Byrne & Dillon, 1986; Cox & Moore, 1988; Seewald & Ross, 1988; Skinner, 1988). Recently, a representation of the LTASS has been recommended for use in amplification research in general and for the calculation of gain in hearing aid prescription formulas in particular (Cox & Moore, 1988).

Traditionally, the LTASS has been measured by recording continuous discourse with a microphone placed directly in front of the talker. When measured at this location the primary energy content of the 1/3-octave band levels of the LTASS (males and females combined) is in the frequency bands centered between 250-630 Hz (Cox & Moore, 1988). The 1/3-octave band levels of the LTASS decreases by approximately 5-6 dB/octave for the frequency bands between 500-4000 Hz (Cox & Moore, 1988). Two factors that have been shown to affect the frequency content of the LTASS are: the level of vocal effort used by the talker and the position of the measurement microphone.

The level of vocal effort used by the talker has an effect on the frequency band levels of the LTASS. Specifically, as vocal effort increases from normal to greater levels the mid-frequency 1/3-octave band levels of the LTASS increase in intensity relatively more than do the low and high frequency 1/3-octave band levels (Brandt, Ruder, & Shipp, 1969; Licklider, Hawley, & Walkling, 1955; Pearsons, Bennett, & Fidell, 1977).

Dunn and Farnsworth (1939) examined the relative changes in the LTASS of one individual when the spectrum was obtained from different positions in space relative to the talker's mouth. The overall level of speech (dB SPL) and the band level in 13 contiguous frequency bands were measured. Repetitions of the speech sample were made while measurements for 76 different microphone placements were obtained. The position of the recording microphone relative to the talker's mouth included various combinations of: distance, azimuth, and altitude. For each microphone position Dunn and Farnsworth (1939) reported the level in each frequency band relative to the level for the same frequency band at the microphone recording position of 30 cm radius, 0° azimuth and 0° altitude. The results indicated that the relative level of the frequency components of the LTASS varied as a function of microphone placement in a sphere around the talker.

Recently, Cornelisse, Gagné, and Seewald (1991) compared the LTASS measured at two different recording positions: 30 cm directly in front of the talker (reference position) and at the tragus of the talker's ear (ear-level position). Speech samples were recorded at both microphone locations simultaneously using two miniature microphones and a dual channel tape recorder. Subsequently, the recordings were analyzed to measure the LTASS of individual subjects. Mean LTASS were calculated for three groups of subjects: adult males, adult females, and children. The results obtained for the three groups were similar. Specifically, significant differences were obtained in the overall level and the frequency spectra between the LTASS at each microphone location. In particular, the overall level of the LTASS measured at the ear-level position was approximately 5 dB greater than the overall level obtained at the reference position. For all three groups of subjects the absolute levels of the LTASS measured at the ear-level position consisted of more low frequency energy (i.e., below 1000 Hz) and less high frequency energy (i.e., above 2500 Hz) than did the LTASS measured at the reference microphone position. The findings suggested that the sensation level of the high frequency energy of a talker's own speech productions, amplified via an ear-level hearing aid, may be different from that predicted by a hearing aid prescription formula that was based upon a typical representation of the LTASS (i.e., measured with a microphone positioned in front of the talker).

The relationship between the input level of a speech signal and the sensation level of the amplified speech has been illustrated previously (Cornelisse, et al., 1991). For the case where the input speech level is less than the level of the LTASS used to calculate gain, the sensation level of the amplified speech signal will be lower than that predicted by the gain algorithm. Conversely, for the case where the input speech level is greater than the level of the LTASS used to calculate gain, the sensation level of the amplified speech signal will be higher than that predicted by the gain algorithm. Therefore, it is important that the representation of the LTASS used to calculate required real-ear gain and to verify real-ear hearing aid performance incorporate the average speech signal received at the microphone of the amplification device. The transfer function for a chest-level microphone could be used to derive a representation of the LTASS obtained via the chest-level microphone of an FM amplification system. Very little data concerning the transfer function for a chest-level microphone are available. Therefore, an experiment designed to determine the extent of differences between the LTASS obtained at the reference microphone position and at the chest-level of a talker was performed.

Method

Subjects

Ten adult females ranging in age between 22 and 28 years (mean = 24 years) participated in the present investigation. All subjects were native speakers of North American English and were able to read fluently from the printed text.

Environment and Apparatus

Each recording of speech was made in a 2.92m x 2.92m (horizontal) x 2.07m (vertical) double-wall sound treated room. Speech recordings were obtained from two different recording positions: directly in front of the talker (reference position) and with a microphone placed on the talker's chest (chest-level position). For recordings obtained at the reference position a microphone was mounted on a tripod 30 cm directly in front of the talker (position of 0° azimuth and 0° altitude). The diaphragm of the probe microphone at the reference position was oriented towards the talker. For recordings obtained at the chest-level position a microphone was attached to a neck-loop and rested on the talker's chest, approximately 20 cm below the mouth (position of approximately -80° altitude). The diaphragm of the probe microphone at the chest-level position was oriented towards the talker's mouth. The chest-level microphone position was chosen to approximate the location of a body worn hearing aid microphone and/or an FM transmitter microphone.

Two Quest miniature probe microphones (Knowles electret condenser microphones — type BT 1759) were used for recording the speech signal. Two Quest preamplifiers (PA-4) served to connect the two microphones to two Quest sound level meters (model 155). The AC output of each sound level meter was connected to one of the input channels of a TEAC dual channel master reel-to-reel tape recorder (model A-7300). The speech signal from each probe microphone was recorded on individual tracks of studio quality audio tape for subsequent analysis. The 1/3-octave band frequency response of each probe microphone and record-playback channel combination was calibrated to within ±0.5 dB. The long-term average ambient noise level in the recording booth did not exceed 24 dB SPL in any 1/3-octave band.

Procedure

Recording of Speech Samples

Prior to the recording session an acoustic calibration tone of approximately 30 seconds in duration was recorded. During each recording session one probe microphone was placed at the reference position and the second probe microphone was placed at the chest-level position. To control for slight differences in the frequency response between the two probe microphones, two speech recordings were obtained at each microphone position (one with each probe microphone).

Each talker was required to read orally a short segment of text entitled "The Tree-house" taken from a grade three reader (McInnes, 1972). The subjects were instructed to talk at a normal rate and with normal vocal effort. The subjects were instructed to disregard any reading errors and to continue reading until the end of the text segment. Each subject was required to read the entire segment of text out loud once for practice. Then, the probe microphones were positioned and the subject was recorded while reading the same segment of text. Finally, the probe microphones at each position were interchanged, and the subject was recorded while reading the same segment of text again.

Acoustic Analysis of Speech Samples

The recorded speech samples were subsequently replayed for spectral analysis. A Bruel and Kjaer dual channel signal analyser (type 2032) with graphic recorder (type 2313) and special application package (type 7066) was used to measure: (1) the long-term average overall level of speech and (2) the long-term average 1/3-octave band levels for center frequencies ranging from 100 to 10,000 Hz. A Hanning window with maximum overlap of each sample ensemble was used during data analysis. The results obtained for individual subjects were based on the first 900 samples of speech (i.e., 2 minutes 10 seconds of recorded speech) averaged by the signal analyser. The mean of the 1/3-octave band levels from both probe microphones was used as the representation of individual LTASS at each microphone position.

Results

Overall Level

The mean overall level of the LTASS measured with the probe microphone placed at the reference position was 65.4 dB SPL (sd=3.4). The mean overall level of the LTASS measured with the probe microphone placed at the chest-level position was 70.8 dB SPL (sd=3.6). The results of an independent t-test analysis revealed that the mean overall level measured at the chest-level position was significantly greater (t=34.780, df=9, p<0.01) than the mean overall level measured at the reference microphone position. This result may be explained in part by the fact that the chest-level microphone was approximately two-thirds the distance of the reference microphone from the sound source (i.e., talker's mouth).

1/3-octave Band Levels

The mean 1/3-octave band levels of the LTASS obtained at the reference microphone position and at the chest-level position are plotted in Figure 1. The 1/3-octave band levels of the ambient noise in the recording environment are also plotted in Figure 1. The results indicated that the 1/3-octave band levels of the LTASS measured at the chest-level position consisted of more low frequency energy (i.e., below 2000 Hz) than did the LTASS measured at the reference microphone position. The results also indicated that the 1/3-octave band levels of the LTASS measured at the chest-level position were similar to the 1/3-octave band levels of the LTASS measured at the reference microphone position for the high frequencies (i.e., above 2000 Hz). A one-way, repeated measures ANOVA of the difference between the 1/3-octave band level at the reference position and the 1/3-octave band level at the chest-level position revealed significant differences (F=27.085, df=20,180, p<0.001) as a function of frequency. The analysis indicated that differences in the frequency spectra observed at the two microphone positions were not constant across frequency. Therefore, the differences between the LTASS obtained at the chest-level position and at the reference microphone position cannot be accounted for solely by the difference in overall level observed at the two microphone positions.

Discussion

Transfer Function of Chest-level LTASS

The transfer function of the LTASS obtained at the chest-level position is plotted in Figure 2. The transfer function was calculated by subtracting the relative 1/3-octave band level (dB re: overall level) of the LTASS obtained at the chest-level position from the relative 1/3-octave band level of the LTASS obtained at the reference microphone position. When calculated in this manner the difference in the overall level of speech obtained at the two microphone positions is not included in the transfer function. That is, the transfer function represents a difference in the 1/3-octave band level resulting from a change in the microphone position (i.e., azimuth and altitude), but does not include a difference that would result from a change in the distance of the reference microphone from the talker (i.e., overall level). A positive transfer function value indicates that the relative 1/3-octave band level of the LTASS obtained at the chest-level position was greater than the relative 1/3-octave band level of the LTASS obtained at the reference position. The transfer function of the LTASS from Dunn and Farnsworth (1939), which approximates the present chest-level position, is included for comparison. The LTASS reported by Dunn and Farnsworth (1939) was obtained from a single adult male subject at a microphone

Figure 1. A plot of the 1/3-octave band levels of the LTASS obtained at the chest-level and reference microphone positions. The data represent the mean of ten adult females.

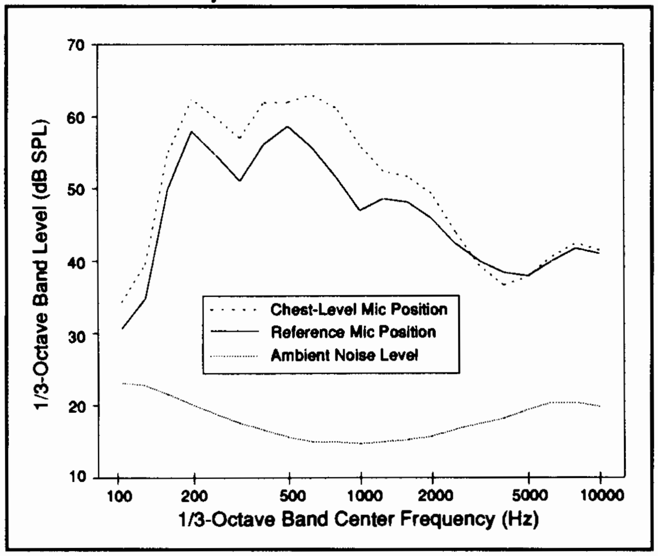
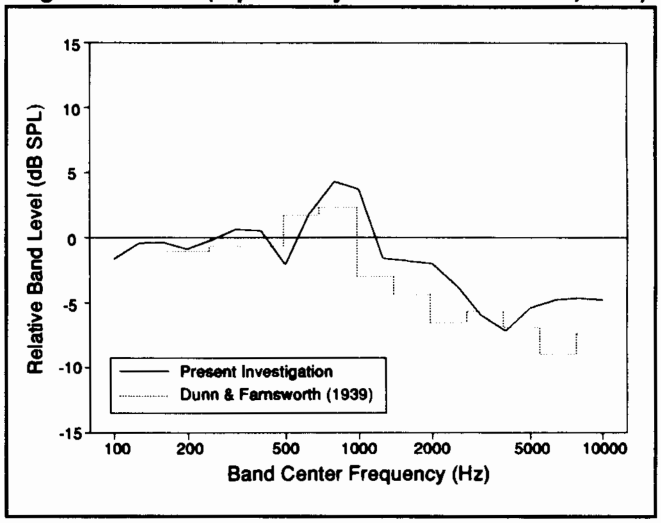


Figure 2. Transfer function of the LTASS obtained at the chest-level microphone position calculated for the group of adult females from the present investigation and a single adult male (reported by Dunn & Farnsworth, 1939).



position of 15 cm distance from the talker's lips (-90° altitude). Both transfer functions are similar and reveal a gradual attenuation of the high frequency components (i.e., above 1600 Hz) of the LTASS obtained at the chest-level position.

LTASS at Chest-level Body Aid Microphone

The chest-level LTASS obtained in the present investigation can be used to estimate the sensation level of a talker's own amplified speech production received via the chest-level microphone of a body aid. The mean 1/3-octave band levels of the LTASS obtained at the chest-level microphone position are plotted in Figure 3 along with the LTASS-HA suggested by Cox and Moore (1988) for use in the calculation of hearing aid gain. The overall level of these two representations of the LTASS is within 1 dB (70.8 and 70.0 dB SPL for the chest-level and Cox and Moore (1988) spectra, respectively). Both representations of the LTASS are generally similar for frequency bands below 2500 Hz. The primary difference between these two representations of the LTASS is for the high frequency bands (i.e., above 2500 Hz) for which the LTASS obtained at the chest-level position differs from the LTASS-HA suggested by Cox and Moore (1988) by as much as 9.5 dB at 4000 Hz. This finding suggests that the mean sensation level of the high frequency energy of a talker's own chest-level amplified speech productions will be less than that predicted by hearing aid gain algorithms.

Derivation of LTASS for FM Systems

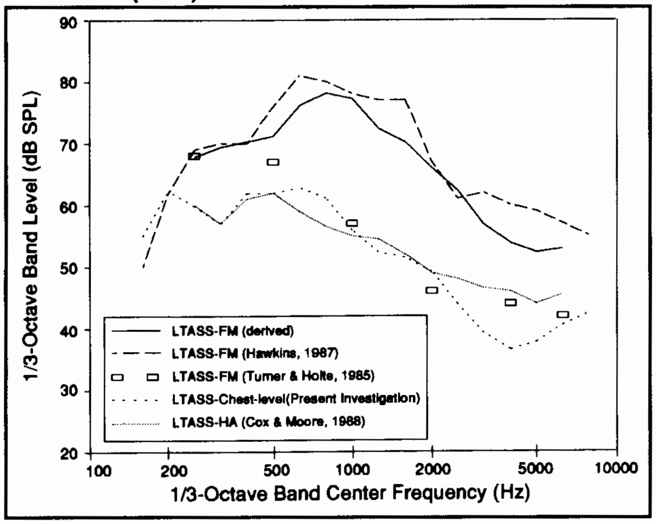
The transfer function of the LTASS obtained at the chest-level position can be used to derive an estimate of the LTASS obtained with an FM system chest-level microphone (LTASS-FM). The LTASS at the microphone of a teacher's FM transmitter will be different from the LTASS for hearing aids proposed by Cox and Moore (1988) for two reasons: (1) teachers generally talk with increased vocal effort (Pearsons et al., 1977); and (2) the microphone of most FM assistive listening devices is worn at chest level of the talker, approximately 20 cm below the mouth (Turner & Holte, 1985).

Data reported by Pearsons et al. (1977) were used to calculate changes in the LTASS resulting from an increase in vocal effort (see Table 1). The correction factor for vocal effort was calculated by subtracting the 1/3-octave band level of the LTASS for "normal" vocal effort from the 1/3-octave band level of the LTASS produced with "loud" vocal effort and represents the average of the data from the group of adult males and the group of adult females. The mean overall level of the LTASS produced by adult male and female talkers using "loud" vocal effort was 74 dB(A) at 1 meter (Pearsons et al., 1977). This level is similar to the average overall level of speech produced by teachers talking in a classroom – 71 dB(A) at a microphone distance of 1 meter (Pearsons et al., 1977).[1]

The data used to derive the LTASS-FM are listed in Table 1. The correction factor for vocal effort (calculated from Pearsons et al., 1977) and the correction factor for the chest-level microphone position (present investigation) were

1 The difference in overall level between "normal" vocal effort and "loud" vocal effort was 14.5 dB.

Figure 3. LTASS-FM derived from the present investigation, LTASS-FM suggested by Hawkins (1987), LTASS-FM suggested by Turner and Holte (1985), chest-level LTASS (present investigation), and LTASS-HA suggested by Cox and Moore (1988).



both added to the LTASS-HA (Cox & Moore, 1988) to obtain the derived LTASS-FM. The derived LTASS-FM is plotted in Figure 3 along with the LTASS-FM suggested by Hawkins (1987) and the LTASS-FM suggested by Turner and Holte (1985).

The derived LTASS-FM is very different from the LTASS-FM suggested by Turner and Holte (1985). The differences between the LTASS-FM suggested by Turner and Holte (1985) and the derived LTASS-FM are due primarily to a difference in the overall level between these two representations of the LTASS. It would appear that the LTASS recommended by Turner and Holte (1985) did not take into account differences in vocal effort typically displayed by teachers. The derived LTASS-FM is similar to the spectrum suggested by Hawkins (1987). The overall level of the derived LTASS-FM is 84.5 dB SPL which is equal to the overall level of the LTASS-FM suggested by Hawkins (1987). The distribution of speech energy for these two representations of the LTASS-FM is also similar suggesting that the primary factors affecting the LTASS received at the chest-level microphone of an FM system amplification device worn by a teacher are: vocal effort and the transfer function for the chest-level microphone position.

Summary

Results of the present investigation indicated significant differences in the overall level and frequency spectra between the mean LTASS measured at the reference and chest-level microphone locations. The primary difference between the LTASS obtained at chest-level versus the representation of the LTASS-HA (Cox & Moore, 1988) used to calculate hearing aid gain is for the high frequency bands (i.e., above 2500 Hz). The finding of reduced high frequency energy for the LTASS obtained at the chest-level position suggests that the mean sensation level of the high frequency energy of a talker's own chest-level amplified (i.e., body aid) speech productions will be less than that predicted by hearing aid gain algorithms. The derived LTASS-FM is similar to the spectrum previously suggested for use in the evaluation of FM systems (i.e., Hawkins, 1987). The representation of the LTASS-FM accounts for spectral changes resulting from increased vocal effort and the chest-level location of the FM microphone.

Table 1. Data required for the calculation of the LTASS-FM (dB SPL). The second column is the speech spectrum for hearing aids (LTASS-HA suggested by Cox & Moore, 1988). The third column is the correction factor for the chest-level microphone (data from present investigation). The fourth column is the correction factor for vocal effort (data from Pearsons et al., 1977). The two correction factors are added to the LTASS-HA to obtain the derived LTASS-FM. See text for explanation of correction factors used.

		Correction for:		
Frequency	LTASS-HA	Chest-level	Vocal-effort	Derived LTASS-FM
250	60.0	-0.2	8.0	67.8
315	57.0	0.6	11.8	69.4
400	61.0	0.5	8.8	70.3
500	62.0	-2.1	11.3	71.2
630	59.0	1.9	15.3	76.1
800	56.5	4.3	17.3	78.1
1000	55.0	3.7	18.5	77.2
1250	54.5	-1.6	19.5	72.4
1600	52.0	-1.8	20.0	70.2
2000	49.0	-2.0	19.0	66.0
2500	48.0	-3.7	18.0	62.3
3150	46.5	-6.0	16.3	56.8
4000	46.0	-7.2	15.0	53.8
5000	44.0	-5.4	13.8	52.3
6300	45.5	-4.8	12.3	52.9

Acknowledgements

The first author (L.E.C) was awarded the CASLPA Annual Student Paper Award (Audiology) based upon a previous version of this manuscript. The research was supported by funds provided by the Academic Development Fund, University of Western Ontario; the Ontario Ministry of Health; the University Research Incentive Fund; and Unitron Industries Limited.

Address all correspondence to: Jean-Pierre Gagné, PhD, (or) Richard C. Seewald, PhD, Hearing Health Care Unit, Department of Communicative Disorders, Elborn College, The University of Western Ontario, London Ontario Canada N6G 1H1

References

Brandt J., Ruder K., & Shipp T. (1969). Vocal loudness and effort in continuous speech. *Journal of the Acoustical Society of America, 46*, 1543-1548.

Byrne D., & Dillon H. (1986). The National Acoustic Laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear and Hearing, 7*, 257-265.

Cornelisse L.E., Gagné J-P., & Seewald R.C. (1991). Ear level recordings of the long-term average spectrum of speech. *Ear and Hearing, 12*, 47-54.

Cox R.M., & Moore J.N. (1988). Composite speech spectrum for hearing aid gain prescriptions. *Journal of Speech and Hearing Research, 31*, 102-107.

Dunn H.K., & Farnsworth D.W. (1939). Exploration of pressure field around the human head during speech. *Journal of the Acoustical Society of America, 10*, 184-199.

Hawkins D.B. (1987). Assessment of FM systems with an ear canal probe tube microphone system. *Ear & Hearing, 8*, 301-303.

Licklider J., Hawley M., & Walking R. (1955). Influences of variations in speech intensity and other factors upon the speech spectrum. *Journal of the Acoustical Society of America, 27*, 207.

McInnes J.A. (Ed.). (1972). *Treats and Treasures.* Don Mills, Ontario, Thomas Nelson & Sons Limited ("The Tree-house", pp. 10-18).

Pearsons K.S., Bennett R.L., & Fidell S. (1977). *Speech levels in various noise environments* (Report No. EPA-600/1-77-025). Washington, DC. 20460. US Environmental Protection Agency.

Seewald R.C., & Ross M. (1988). Amplification for young hearing impaired children. In M.C. Pollack (Ed.), *Amplification for the Hearing Impaired.* New York, NY, Grune and Stratton.

Skinner M.A., (1988). *Hearing Aid Evaluation.* Englewood Cliffs, NJ, Prentice Hall.

Turner C.W., & Holte L.A., (1985). Evaluation of FM amplification systems. *Hearing Instruments, 36(7)*, 9-12,56.