Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Ecological loudness: Binarual loudness constancy International Congress on Acoustics, ICA 2010

Mary Florentine (1) and Michael Epstein (2)

(1) Communication Research Laboratory, Dept. of Speech-Language Pathology and Audiology, and Institute for Hearing, Speech & Language, Northeastern University, 360 Huntington Ave - 106A Forsyth Building, Boston, MA 02115, United States of America
(2) Auditory Modeling and Processing Laboratory, Dept. of Speech-Language Pathology and Audiology, and CDSP Center, Dept. of Electrical and Computer Engineering, Northeastern University, 360 Huntington Ave, Boston, MA 02115, United States of America

PACS: 43.66.CB, 43.66.PN, 43.66.RQ

ABSTRACT

Are conclusions about loudness drawn from tones presented via earphones in laboratories applicable to listening to a talker in a room? The present experiment tests the following hypothesis: speech from the same talkers presented under more ecologically valid conditions results in a smaller binaural-to-monaural loudness ratio than speech presented without visual cues and/or presented via headphones. Twelve normal listeners were presented two types of stimuli (recorded speech, with and without visual cues) monaurally and binaurally across a wide range of levels. The same stimuli were presented via earphones and loudspeakers. Loudness was measured using magnitude estimation. Results show that the binaural-to-monaural loudness ratio was significantly less for speech with visual cues presented via a loudspeaker than for stimuli with any other combination of test parameters (i.e., speech without visual cues presented via both headphones and loudspeakers, and speech presented with visual cues via headphones). The present results indicate that the loudness of a visually present talker in daily environments is little affected by switching between binaural and monaural listening. This phenomenon has been dubbed “Binaural Loudness Constancy,” because of its similarity to loudness constancy that occurs with distance from the speaker. The present experiment supports the importance of ecological validity in loudness research, which could change how perception of loudness is understood.

INTRODUCTION

Laboratory experiments using earphones indicate that a tone presented binaurally is louder than the same tone presented monaurally (Fletcher and Munson, 1933). It has been generally assumed that the binaural-to-monaural loudness ratio is equal to two for dichotic tones at the same loudness [i.e., a tone presented to two ears is twice as loud as a tone presented to only one ear. (For review see, Sivonen and Ellermeier, 2010)]. More recent studies suggest a lower ratio ranging from 1.29 to 1.7 (Epstein and Florentine, 2009; Marozeau, Epstein, Florentine and Daley, 2006; Scharf and Fishken, 1970; Zwicker and Zwicker, 1991). When real sound sources are heard in an anechoic chamber, an average binaural-to-monaural loudness ratio of 1.22 has been reported (Sivonen, 2006).

New data obtained using speech stimuli from a visually present talker heard via loudspeakers challenge conclusions drawn from classical binaural measurements obtained in laboratories using earphones. Epstein and Florentine (2009) presented preliminary data showing that (1) the binaural-to-monaural loudness ratio is significantly smaller for speech from a visually present talker than for recorded speech and tones, (2) the binaural-to-monaural loudness ratio is significantly smaller for loudspeaker presentation than for earphone presentation, and (3) the binaural-to-monaural loudness ratio is smallest for speech from a visually present talker presented via loudspeakers than any of their other test conditions.

Because Epstein and Florentine’s (2009) experimental design was an initial approach to studying this complex problem, it was not feasible to control for differences among talkers. The present experiment builds on their earlier findings and tests the following hypothesis: speech from the *same* talkers presented under more ecologically valid conditions results in a smaller binaural-to-monaural loudness ratio than speech presented without visual cues and/or presented via headphones. To provide a condition that had more ecological validity—while being experimentally defined—loudspeaker presentation of speech with visual cues was chosen. This condition was compared to the same stimuli presented in three other conditions: without visual cues presented via a loudspeaker, without visual cues presented via earphones, and with visual cues presented via earphones.

METHOD

Stimuli

A set of 16 speech stimuli was recorded with video and audio consisting of four native talkers (two males, two females) speaking two-syllable words that have equal stress on both syllables (i.e., spondees). Each of the speech stimuli was presented at five levels: 25, 40, 55, 70, and 80 dB SPL.

Procedure and apparatus

Each listener estimated loudness by assigning a number whose magnitude matched the loudness of each stimulus. Trials were split into eight blocks. Each block contained a permutation of three parameters: video and audio or audio only, presentation via earphones or loudspeaker, and presentation monaurally or binaurally. Each of these eight blocks was presented three times resulting in 24 total blocks. These 24 blocks were presented to the listener in random order. Each presentation block contained 80 trials: four words spoken by four talkers at five levels, presented in random order. This resulted in 1920 total trials per listener.

For the "monaural" loudspeaker condition, a foam earplug (E.A.R. classic - 29 NRR) was inserted. Earplugs increased threshold by 20-24 dB at 1 kHz compared to unoccluded ears, as determined by previous measurements (see Epstein and Florentine, 2009). Although it is impossible to obtain a completely monaural condition in listeners with normal bilateral hearing, the attenuation provided by a monaural earplug was deemed to be sufficient for the present experiment based on data obtained by Zwicker and Zwicker (1991).

The stimuli were presented in MATLAB (2007b running on Windows XP) and converted from digital (48-kHz sampling frequency) to analog using a 32-bit Lynx Two Soundcard. The analog signal was then presented either via Sony MDR-V6 headphones or a single M-Audio BX8a 8-inch studio monitor speaker, 1 m directly in front of the listener in a room that was approximately 2.85 m by 3.07 m. The listener's head was positioned against a headrest that contacted only the back of the head while the listener sat in a straight-backed chair. Levels were calibrated at the approximate location of a listener's ear. Video was displayed on a 17" flat-panel monitor. Listeners typed in magnitude estimates for each stimulus presentation and then the screen was cleared so that no prior estimates were visible to the listener.

Subjects

The subjects consisted of 12 listeners (6 males, 6 females), aged 20-33 years. All had clinically normal hearing.

RESULTS

The binaural-to-monaural loudness ratio was computed using the geometric mean of the binaural magnitude estimates for all listeners and the geometric mean of the monaural magnitude estimates for all listeners. The ratios are: 1.34 for video and audio presented via earphones, 1.29 for audio only presented via earphones, 1.22 for audio only presented via a loudspeaker, and 1.09 for video and audio presented via a loudspeaker. The results clearly show that the condition with the video and audio presented via a loudspeaker has a smaller binaural-to-monaural loudness ratio than any of the other conditions. This observation is supported by statistical analysis. A repeated-measures General Linear Model was applied to the data (SPSS 17.0). The logs of the geometric means of the three magnitude estimates were used. The cross factor of transducer type (earphones or loudspeaker) X presentation mode (video and audio or audio only) X number of ears (one or two) was significant ($p=0.05$; $df=1$; $F=4.74$).

DISCUSSION

The present results lend support to the proposed hypothesis that speech presented under conditions that are closer to typical daily environments result in a smaller binaural-to monaural ratio than speech presented under less ecologically valid conditions. Specifically, the binaural-to-monaural loudness ratio is significantly smaller for spondees with video and audio presentation via loudspeakers than any of the other three conditions. These data are in reasonable quantitative agreement with data from the literature. Most notably, the binaural-to-monaural loudness ratio for the video and audio presentation of speech via a loudspeaker in the present experiment is 1.09 and the mean loudness ratio for a similar condition in Epstein and Florentine's (2009) experiment was 1.08. The binaural-to-monaural loudness ratio across both conditions (audio, audio and video) for earphone presentation in the present experiment is 1.32, which is within the range of most other studies referenced in the introduction (i.e., 1.29 to 1.7).

Taken together, the available data support the concept of Binaural Loudness Constancy in which speech (perhaps all natural sounds) from a visually present source heard under ecologically valid conditions is not much louder when listening with two ears than when listening with only one ear (see Epstein and Florentine, 2009). Although the present experimental design did not permit control of all variables, which need to be examined in detail in future studies, the present study offers further insight into this question. Note that this phenomenon is not only due to the visual cues; the data from audio and video presentation via earphones did not result in a significantly smaller loudness ratio than the same stimuli presented without video.

In daily environments, there are a multitude of cues that a listener may use. In survival situations, it could be advantageous if a listener were able to accurately judge the distance of an environmentally relevant sound whether listening with one or two ears. The results of the present experiment are consistent with the importance of ecological validity in loudness research, which could change how perception of loudness is understood. In conclusion, the present results are intriguing and warrant further study.

ACKNOWLEDGMENTS

Kevin Reilly provided assistance and equipment for stimulus recordings. Tao Cui, Eliza Floyd, Abigail Seaser, Yui Anzai, and Maria Zilberberg assisted with data acquisition and analysis. Whitney Danse, Amy Levasseur, Thomas DiCicco, and Michael Epstein served as talkers.

REFERENCES

Epstein, M. and Florentine, M. "Binaural loudness summation for speech and tones presented via earphones and loudspeakers," *Ear and Hearing* 30(2), 234-237 (2009)

Fletcher, H., and Munson, W.A. "Loudness, its definition, measurement and calculation," *J. Acoust. Soc. Am.* 5, 82-108 (1933)

Marozeau, J., Epstein, M., Florentine, M., and Daley, B. "A test of the binaural equal-loudness-ratio hypothesis for tones," *J. Acoust. Soc. Am.* 120, 3870-3877 (2006)

Scharf, B., and Fishken, D. "Binaural summation of loudness: reconsidered," *J. Exp. Psych.* 86, 374-379 (1970)

Sivonen, V.N.P.E., W. "Directional loudness in an anechoic sound field, head-related transfer functions, and loudness summation," *J. Acoust. Soc. Am.* 119, 2965-2980 (2006)

Sivonen, V.N.P.E., W. and Ellermeier, W. "Binaural loudness," in Loudness ed. M. Florentine, R. Fay, & A. Popper, (Springer, New York, 2010) in press

Zahorik, P., and Wightman, F.L. "Loudness constancy with varying sound source distance," *Nat. Neurosci.* 4, 78-83 (2001)

Zwicker, E., and Zwicker, U.T. "Dependence of binaural loudness summation on interaural level differences, spectral distribution, and temporal distribution," *J. Acoust. Soc. Am.* 89, 756-764 (1991)