Spatial hearing advantages in everyday environments

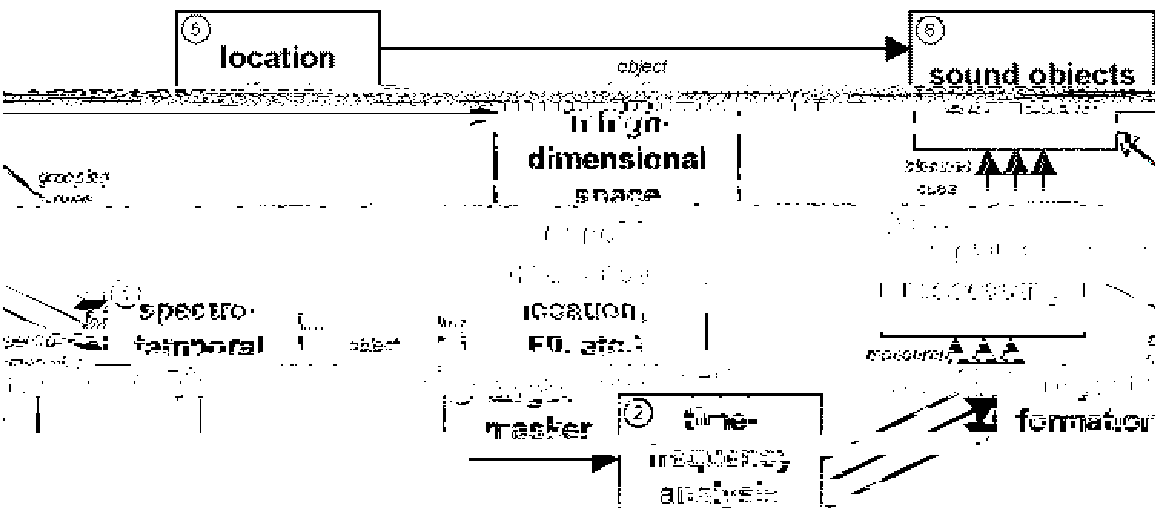
Barbara G. Shinn-Cunningham
Boston University Hearing Research Center
Departments of Cognitive and Neural Systems and Biomedical Engineering
Boston University, Boston, MA 02215
617-353-5764
shinn@cns.bu.edu

Spatial auditory cues are not a dominant factor in human auditory scene analysis (i.e., in "parsing" the sound reaching the ears to determine the number and spectral content of competing sound sources) [1-5]. However, spatial hearing is very important for *understanding* a target source in an environment that has multiple sound sources [6-9]. Resolving this apparent paradox is critical for understanding how human listeners operate in difficult conditions, for instance when there is a heavy workload and there are competing demands on attention. Such knowledge is very important for designing effective auditory displays and other human-machine interfaces.

Traditional views of the benefits of spatial hearing [10] fail to explain this contradiction as well as other observations, such as the 1) relatively poor ability of hearing impaired listeners to parse and understand speech in situations with competing sources and/or reverberation [11], 2) relatively large inter-subject differences in performance on tasks involving "informational masking" compared to tasks in which the masker is dissimilar from the target [11-13], and 3) very large improvements in speech intelligibility that can arise when similar competing talkers arise from different locations compared to when they are in the same location [6-9, 14, 15]. This short paper provides a preliminary conceptual framework that unifies these seemingly contradictory results by isolating and identifying multiple ways in which spatial hearing impacts the ability to listen to competing, simultaneous sound sources.

It has long been known that spatial separation of a target from an interfering source (a masker) improves a listener's ability to detect and understand the content of the target (a phenomenon known as "spatial unmasking;" e.g., see [16-18]; recent reviews include [10, 19]). Much of this improvement can be attributed to simple acoustic effects: spatially separating the target and masker generally increases the target-to-masker energy ratio (TMR) at one of the two ears. Because speech intelligibility improves with TMR, the improvement in TMR in one ear leads directly to an improvement in performance. The acoustic TMR varies with frequency because the acoustic interaction of the head and body of the listener with an impinging sound wave varies with the sound wavelength. Thus, the TMR changes more with spatial location of target and masker at high frequencies than at low frequencies. In the most extreme cases (i.e., when one of the sources is attending to these components in the face of the masker competition. Under these circumstances, perceived differences in the locations of target and masker can improve the ability to attend to and understand the target message [6-9, 11, 23, 24]. In fact, in these situations, spatial separation can cause extraordinarily large changes in performance that even exceed the effects that have been the focus of traditional study (i.e., changes in the acoustic TMR and the effective TMR). Recent work suggests that this form of spatial hearing advantage arises because a listener can selectively attend to a source from a particular location and ignore a competing source from a different direction.

One final factor that has a large impact on spatial hearing advantages in typical environments is the effect of room acoustics [21, 25-29]. Echoes and reverberation alter nearly all acoustics aspects of the signals reaching the ears (spectro-temporal properties, TMR, binaural cues). As a result, the importance of spatial hearing for spatial unmasking differs in “everyday” environments compared to the anechoic conditions under which many psychophysical tests have been performed. Reverberation and echoes degrade spatial unmasking advantages for many traditional test conditions, but are less detrimental on tasks that involve similar-quality, simultaneous talkers and spatial attention. Further, while existing binaural processing models predict spatial unmasking under certain circumstances (e.g., a speech target in the presence of a steady-state, noise masker in anechoic space), they cannot account for the effects of reverberation on speech intelligibility or spatial unmasking.



In order to gain insight into how spatial hearing influences task performance in everyday listening conditions, it is helpful to consider how sound is processed in the auditory system. The spatial auditory pathway is organized in a very hierarchical manner. Even a cursory consideration of its structure suggests that spatial hearing may influence auditory processing at many different stages of processing, and that the importance of spatial hearing for a particular task depends on the nature of the stimuli being presented and the task being performed. The figure above presents our conceptual framework for understanding how spatial hearing influences the ability to listen to a target in the presence of a masking source.

Spatial separation first influences the acoustic TMR through physical interactions, external to the listener (1). Acoustic information is then analyzed neurally to extract spectro-temporal content in the monaural signals reaching the ears (2). This information is processed binaurally in the brainstem (3). We hypothesize that this low-level binaural processing provides information to two parallel processing stages. In one stage, spectro-temporal content of the signals reaching the listener are grouped into acoustic objects by combining spectro-temporal features of the signals reaching the ears (4). We believe binaural processing contributes to this process by revealing spectro-temporal features of a masked signal that may not be audible in a monaural representation (3)-(4). In the grouping stage (4), cues such as harmonicity, common onset, and other features determine how auditory objects are formed. The resulting grouping rules in turn

[8] D. S. Brungart and B. D. Simpson, "Within-ear and across-ear interference in a cocktail-party listening task," *J Acoust Soc Am*, vol. 112, pp. 2985-2995, 2002.
[9] D. S. Brungart and B. D. Simpson, "The effects of spatial separation in distance on the informational and energetic masking of a nearby speech signal," *J Acoust Soc Am*, vol. 112, pp. 664-676, 2002.
[10] P. M. Zurek, "Binaural advantages and directional effects in speech intelligibility," in *Acoustical Factors Affecting Hearing Aid Performance*, G. Studebaker and I. Hochberg, Eds. Boston, MA: College-Hill Press, 1993.
[11] T. L. Arbogast, "The effect of spatial separation on informational and energetic masking of speech in normal-hearing and hearing-impaired listeners," unpublished Ph.D. dissertation, *Dept Comm Disorders*, Boston University, 2003.
[12] E. L. Oh and R. A. Lutfi, "Nonmonontonicity of informational masking," *J Acoust Soc Am*, vol. 104, pp. 3489-3499, 1998.
[13] D. L. Neff and T. M. Dethlefs, "Individual differences in simultaneous masking with random-frequency, multicomponent maskers," *J Acoust Soc Am*, vol. 98, pp. 125-134, 1995.
[14] G. Kidd, C. R. Mason, and T. Rohtla, "Binaural advantage for sound pattern identification," *J Acoust Soc Am*, vol. 98, pp. 1977-1986, 1995.
[15] J. Kidd, Gerald, C. R. Mason, T. L. Rohtla, and P. S. Deliwala, "Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns," *J Acoust Soc Am*, vol. 104, pp. 422-431, 1998.
[16] A. W. Bronkhorst, "The cocktail party effect: Research and applications," *J Acoust Soc Am*, vol. 105, pp. 1150, 1999.
[17] R. Drullman and A. W. Bronkhorst, "Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation," *J Acoust Soc Am*, vol. 107, pp. 2224-2235, 2000.
[18] A. W. Bronkhorst and R. Plomp, "The effect of head-induced interaural time and level differences on speech intelligibility in noise," *J Acoust Soc Am*, vol. 83, pp. 1508-1516, 1988.
[19] A. W. Bronkhorst, "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acustica*, vol. 86, pp. 117-128, 2000.
[20] B. G. Shinn-Cunningham, J. Schickler, N. Kopco, and R. Y. Litovsky, "Spatial unmasking of nearby speech sources in a simulated anechoic environment," *J Acoust Soc Am*, vol. 110, pp. 1118-1129, 2001.
[21] B. G. Shinn-Cunningham, "Speech intelligibility, spatial unmasking, and realism in reverberant spatial auditory displays," Proc Int Conf Aud Display, pp. 183-186, 2002.
[22] M.L. Hawley, R.Y. Litovsky, and J. Culling, "The "cocktail party problem" with four types of maskers: Speech, time-reversed speech, speech-shaped noise, modulated speech-shaped noise," Proc Assoc Res Otolaryng, p. 161, 2000.
[23] N. I. Durlach, C. R. Mason, G. Kidd, T. L. Arbogast, H. S. Colburn, and B. G. Shinn-Cunningham, "Note on informational masking," *J Acoust Soc Am*, in press.
[24] J. Kidd, Gerald, C. R. Mason, T. L. Arbogast, D. S. Brungart, and B. D. Simpson, "Informational masking caused by contralateral stimulation," *J Acoust Soc Am*, vol. in press, 2003.
[25] B. G. Shinn-Cunningham and N. Kopco, "Effects of reverberation on spatial auditory performance and spatial auditory cues," *J Acoust Soc Am*, vol. 111, pp. 2440, 2002.
[26] B. G. Shinn-Cunningham, S. Constant, and N. Kopco, "Spatial unmasking of speech in simulated anechoic and reverberant rooms," Proc Assoc Res Otolaryng, 2002.
[27] J. F. Culling, H. S. Colburn, and M. Spurchise, "Interaural correlation sensitivity," *J Acoust Soc Am*, vol. 110, pp. 1020-1029, 2001.
[28] C. J. Darwin and R. W. Hukin, "Effects of reverberation on spatial, prosodic, and vocal-tract size cues to selective attention," *J Acoust Soc Am*, vol. 108, pp. 335-342, 2000.
[29] J. F. Culling, Q. Summerfield, and D. H. Marshall, "Effects of simulated reverberation on the use of binaural cues and fundamental-frequency differences for separating concurrent vowels," *Speech Comm*, vol. 14, pp. 71-95, 1994.