VOLUME 11, ISSUE 4, SPRING 2009

Journal of Military and Strategic Studies

Auditory Situation Awareness in Urban Operations

Angelique A Scharine, Tomasz R Letowski, James B Sampson

Introduction

The complexity of the urban environment makes military operations difficult. Soldiers rely on vision to conduct patrols and clear buildings. However, visual cues are frequently obscured or blocked by buildings and civilians. During fighting, smoke and flying debris create further confusion and distraction. In such situations, sound becomes the first, and often the only, source of information about the presence of an enemy or the direction of incoming weapon fire. Even when visual cues are available, auditory cues play a critical role in human orientation and navigation because hearing offers immediate access to a full 360º range of sensory information.

It has been reported that both veterans of urban warfare and soldiers training in urban operations (UO) have difficulties hearing the environment. These difficulties are primarily due, not to failure to detect sound, but to problems locating or recognizing that sound. For example, a soldier might not be able to determine exactly where weapon fire is coming from although he clearly heard it. A familiar sound might not be recognizable because the sound occurred on the other side of a building. Even worse, attention can be misdirected to nearby events and a sound can be attributed to the wrong source. A car's backfire can sound like a gunshot, or a civilian shouting can sound threatening. When this happens, a soldier can make innocent, yet fatal, mistakes.

The attacking and defending forces are both subject to the same acoustic conditions that lead to increased ambiguity in auditory information. Although the defending forces have the advantages of knowing the structures in the urban terrain,

ISSN : 1488-559X

they risk isolation and entrapment if communication with allies is lost. In contrast, the attacking forces tasked with locating weapons and the enemy forces are operating in an unfamiliar environment. They are dependent on the sounds of activity inside buildings to determine if they are occupied.

A number of sensors and covert listening devices have been developed to detect activity within buildings and rooms prior to entry. However, these technologies require set-up and pre-planning, limiting their usefulness to soldiers on the move. Moreover, such aids tend to add to the soldier's cognitive workload, making unaided perceptual skills preferable for the majority of situations.

The purpose of this report is to describe the known effects of the urban acoustic environment on the soldier's ability to locate target sounds and the impact of this on battlefield situation awareness. Although much of this information is based on research that has been conducted under controlled laboratory conditions designed to minimize confounding factors, we apply this research to the operating conditions experienced by the soldier in order to better understand its impact. Cognitive factors, such as attention and experience, often ignored when examining auditory perception, will be highlighted as well. These factors play a documented role in spatial orientation.

Sound Localization Basics

There are a number of cues used for auditory spatial orientation. The relative usefulness of these cues depends on the experience of the listener, the type of environment, and the nature of the sound sources. These cues allow the person to assess the location of the sound source in the horizontal plane (azimuth), location of the sound source in the vertical plane (elevation), the distance to the sound source, and the volume of the acoustic space. In order to discuss how UO conditions may affect a Soldier's ability to localize a sound source in space, let us define these four elements of spatial orientation.

1. Azimuth: The horizontal angle (0 - 359 degrees) from a reference point. Zero is usually defined as straight ahead of the listener. The reference horizontal plane intersects the ears of the listener.

2. Elevation: The vertical angle (0 – 359 degrees) from a reference point. As with azimuth, zero is defined as straight ahead of the listener at the level of the ears of the listener. The reference vertical plane is midway between the ears.
3. Distance: The linear distance of the sound source from the listener.
4. Volume: The size of the surrounding environment manifesting itself to the listener by the sound reverberation that occurs in that space.

Good auditory situation awareness requires that the soldier perceive these dimensions accurately. The human auditory system achieves this by using both binaural and monaural cues. Binaural cues are the dynamic differences between sounds entering the two ears of the listener. Monaural cues come from the reflections of sounds from the listener's outer ear, head, and shoulders. Other cues include head movements, time needed for sound to decay (space reverberation), and the listener's familiarity with the sound sources and the environment. Volume affects localization by informing the listener about the size of the acoustical environment, providing a frame of reference for the calibration of distance cues. The following sections provide information about the specific acoustic cues used to locate sound sources in the horizontal, vertical, and distance dimensions.

Horizontal – Azimuth: The influence of two ears

In Figure 1, the bird's chirping is more intense or louder in the ear that is closer to the sound (t1) because of the "baffling effect" of the head which casts an "acoustic shadow" on the sound coming to the more distant ear (t2). The difference in sound levels from this shadow is called the interaural **level** difference. At low frequencies[1], interaural level differences are very small because the dimensions of the human head and the area of acoustic shadow are small in comparison to the long wavelength of

[1] The terms *frequency* and *spectrum* are referred to often in this paper. A pure tone of a particular frequency has a certain *pitch*, so frequency and pitch are sometimes used as synonyms. However, most naturally occurring sounds are not pure tones, but instead are composed of many frequency components. A description of the frequency components contained in a particular sound and their respective strength or *amplitude* is called that sound's *spectrum*.

sound. At high frequencies, however, interaural level differences are sufficient to provide clear localization cues. The difference in sound intensity can be as much as 8 decibels (dB)[2]. Thus, interaural level differences are powerful localization cues at higher frequencies but they fail at low frequencies.

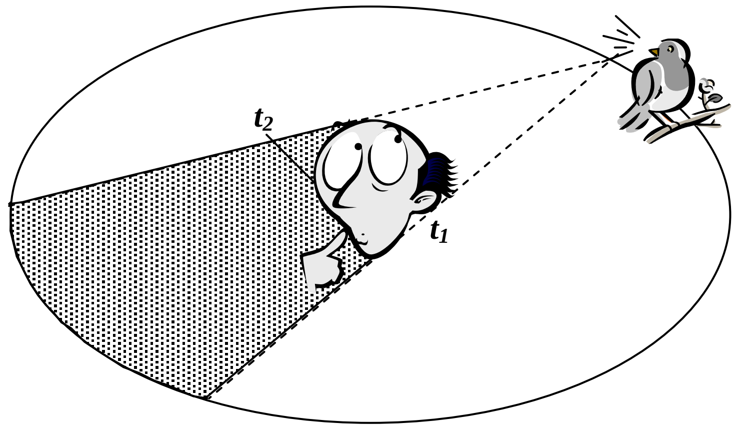

Figure 1. Illustration of interaural time and level differences due to head shadow. The bird's chirps arrive earlier at and are louder at the listener's left ear (t_1) than his right (t_2).

However, the same sound will arrive at the near ear earlier than the far ear and the resulting phase (time) difference helps to localize low frequency sounds. Therefore, at low frequencies sound localization in the horizontal plane depends predominantly on the interaural **time** differences. The time difference is not an effective cue at high frequencies because various time delays may result in the same phase difference at the ears.

[2] J. C. Steinberg and W. B. Snow, "Physical Factors," *Bell Systems Technical Journal* 132 (1934): pp. 245-58.

The two mechanisms described above are the foundation of the "duplex theory" of sound localization.[3] According to this theory, sound source location in space is defined by time differences at low frequencies and sound level differences at high frequencies. Because the frequency ranges in which these binaural cues operate do not overlap, localization errors in the horizontal plane are the largest for sound sources emitting signals in the mid-frequency (1000 Hz to 3000 Hz) range. This effect is exacerbated by the fact that people are very sensitive to sounds in this frequency range and, therefore, sounds in this frequency range constitute added "noise" that is difficult to localize and yet prominent, masking other sounds in the auditory scene.

The "precedence effect"[4] refers to the way the auditory system suppresses echoes and inhibits the effects of reflections, allowing one to localize sounds based on the direct signal. It makes us immune to the effects of small amounts of reverberation. However, this inhibition is limited to a small time window and breaks down if the reflected sound arrives very late, with respect to the direct sound. Strong later reflections tend to "pull" the perceived sound location away from that of the source of the direct sound towards the location of the reflective surface (the last sound heard). The spectral characteristics of these reflections can affect the degree to which they are inhibited and also change the perceived size, loudness, and timbre of the original sound.[5]

Both interaural sound level and time differences provide reliable information about a sound source's position on the left-right axis; however, they do not distinguish between positions in the front and back or at different elevations. Human ability to localize sounds along these dimensions requires the presence of monaural cues.

---

[3] J. W. Strutt, Lord Rayleigh, 3rd Baron of Rayleigh, "On Our Perception of Sound Direction," *Philosophical Magazine* 13 (1907): pp. 214-32.

[4] H. Wallach, E. B. Newman, and M. R. Rosenzweig, "The Precedence Effect in Sound Localization," *The American Journal of Psychology* 62 (1949): pp. 315-36.

[5] P. Divenyi, and J. Blauert, "On Creating a Precedent for Binaural Patterns: When Is an Echo an Echo?" In *Auditory Processing of Complex Sounds*, ed. W. A. Yost and C. Watson (New Jersey: Erlbaum, 1987); J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (Cambridge, MA: MIT Press, 1999).

Elevation

Sound source position along the front-back axis and elevation are determined primarily by monaural cues. Despite the general success of binaural cues and the "duplex theory" in explaining localization of sound sources in space, they are unresolved within a region known among auditory researchers as the "cone of confusion": a cone extending outward from each ear and centered on the lateral axis connecting the two ears of the listener. All locations on this cone have the same binaural differentials (see Figure 2) and cannot be resolved by binaural cues.[6] Therefore, an additional perceptual mechanism is needed to specify the sound source's location on the cone. This is the domain of monaural cues. Monaural cues are directionally dependent spectral changes that take place when sound is reflected from the folds of the ear (pinnae) and the shoulders of the listener. The combined directionally dependent monaural and binaural cues are called the head-related transfer function (HRTF). The resulting spectral changes are largest in the frequency ranges above 4 kHz, approximately, and can be best interpreted in reference to the spectral content of the original sound. The richer in high frequency content the sound is, the more useful the monaural information will be.

[6] S. R. Oldfield, and S. P. A. Parker, "Acuity of Sound Localisation: A Topography of Auditory Space. I. Normal Hearing Conditions," *Perception* 13 (1984): pp. 581-600. This is not strictly true. The "cone of confusion" model assumes a spherical head. However, auditory localization error patterns generally support that this model approximates human behavior well.

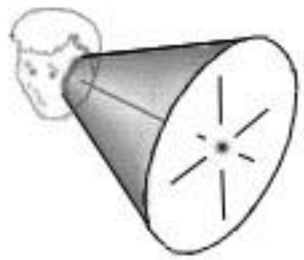

Figure 2. Cone of Confusion. Binaural cues specify a subset of locations that can be described by a cone shaped region. In order to specify one of these possible locations, monaural cues are needed.

People can localize sound sources in the horizontal plane with one ear but the resulting localization error is much greater (~30º-40º) than when binaural cues are available (~3º-4º). Similarly, elimination of monaural cues leads to front-back reversals and errors in elevation. Thus, monaural and binaural cues enhance each other and are required for precise sound localization.

Monaural cue changes occur relative to the original sound source and, therefore, their interpretation requires some familiarity with the original sound source. Yet, we localize rather well much of the time despite the fact that two instances of a sound, familiar or otherwise, are rarely the same. This is because head movements provide additional information to resolve binaural ambiguity. If the sound is long enough (> 0.5s) to allow for sufficient head movement before the sound ceases, the movements aid in sound localization by providing the listener with localization cues at different head positions.[7] For example, if a listener is confusing the front with the back, turning the head to the left or right will cause the sound image to move opposite to the listener's expectations and this feedback will remove the confusion.

[7] W. Noble, "Auditory Localization in the Vertical Plane: Accuracy and Constraint on Bodily Movement," *Journal of the Acoustical Society of America* 82 (1987): pp. 1631-36.

Distance

Three cues are used to estimate auditory distance: intensity (loudness), sound spectrum, and temporal offset (decay). All of these cues require previous knowledge about the original sound source and the acoustical characteristics of the environment. In addition, many people cannot estimate perceived distance because they lack the frame of reference required to give numerical values even when estimating distance visually. Due to the complexity of conditions affecting auditory distance judgments, these judgments are quite inaccurate and result in about 20per cent error, more or less.[8]

The most natural auditory distance estimation cue seems to be sound intensity (loudness)[9]. Under ideal conditions, sound intensity decreases by 6 dB per doubling of the distance from the receiver. Therefore, a comparison of the perceived intensity to the expected intensity of a sound can allow one to estimate the distance. However, although movement toward or away from the operating source may provide a frame of reference[10] use of intensity as a cue requires that the listener know how loud the original sound was.

The actual relationship of attenuation to distance is variable. Changes in the propagating medium (e.g., air), such as temperature and relative humidity, affect the absorption of sound across distance differently for different spectral frequencies. If one is familiar with the environment and its sounds, these changes provide useful distance information.[11] However, without familiarity, changes in intensity provide only relative information about distance.[12]

In rooms and other closed spaces, the decrease in sound intensity may initially follow the 6 dB doubling rule given above but soon become less due to room reflections from nearby surfaces (e.g., the floor) as the distance increases. At greater distances, the

---

[8] B. C. J. Moore, *An Introduction to the Psychology of Hearing* 4th ed. (San Diego, CA: Academic Press, 1989).

[9] D. H. Mershon, and L. E. King, "Intensity and Reverberation as Factors in Auditory-Perception of Egocentric Distance," *Perception & Psychophysics* 18 (1975): pp. 409-15.

[10] D. H. Ashmead, D. LeRoy, and R. D. Odom, "Perception of the Relative Distances of Nearby Sound Sources," *Perception & Psychophysics* 47 (1990): pp. 326-31.

[11] P. McGregor, A. G. Horn, and M. A. Todd, "Are Familiar Sounds Ranged More Accurately?" *Perceptual and Motor Skills* 61 (1985): p. 1082.

[12] A. D. Little, , D. H. Mershon, and P. H. Cox, "Spectral Content as a Cue to Perceived Auditory Distance," *Perception* 21 (1992): pp. 405-16.

decrease in intensity continues as long as the energy of the direct sound exceeds that of the reflected sounds. The distance from a sound source to the point where both sound energies (direct and reflective) are equal is called the area of *critical distance*. Inside the critical distance area, sound localization is mostly unaffected by sound reflections from space boundaries due to the precedence effect. Beyond the critical distance, however, the precedence effect may not operate. Therefore, the listener's estimates of distance will become less accurate the farther away the listener is from the sound source and the more reflective surfaces in the space.

Reverberation is sound reflected from the ground, walls, and other objects. It lasts longer and decays slower than the original sound. As distance between the sound source and the listener increases, the amount of direct sound energy arriving to the listener's ears decreases and the amount of reverberant (reflected) energy increases.[13] Therefore, perceived reverberation constitutes a very effective cue for distance estimation. However, the specific ratio of these two energies is also dependent on the directivity of the sound source, the listener's hearing, the size of the space, sound frequency, and the position of the sound source relative to the walls and the listener. This means that small, highly-reflective spaces may create the same perceptual effects as larger, more damped spaces. Thus, reverberation information coming from unknown and unseen spaces, such as adjacent rooms or buildings, is unlikely to provide usable distance information until the listener becomes familiar with the space.

Auditory Cues and Situation Awareness in the Urban Battlefield

Sound can be the first warning of events occurring around us. Unlike vision, which can be obscured by buildings and people, sound is perceptible from the entire 360° range and can travel around objects to some degree. Thus, auditory information is critical for situation awareness during UO, both as a complement and a supplement to vision.

---

[13] D. H. Mershon, W. L. Ballenger, A. D. Little, P. L. McMurtry, and J. L. Buchanan, "Effects of Room Reflectance and Background-Noise on Perceived Auditory Distance," *Perception* 18 (1989): pp. 403-16; S. H. Nielsen, "Auditory Distance Perception in Different Rooms," *Journal of the Audio Engineering Society* 41 (1993): pp. 755-70.

However, the physical acoustics and the psychological conditions faced by the soldier during UO are not always favourable for contributing to good situation awareness. The localization cues, described above, suggest that accurate sound localization requires a sound source emitting long bursts of broadband spectral energy in an environment with few reflective surfaces and minimal ambient noise. The ideal listener is mobile, undistracted, and familiar with the typical noises in the environment. These conditions are often not met in the urban battlefield.

Acoustic conditions

The physical structures in an urban environment reflect sound. Multiple surfaces create multiple reflections or reverberation. Reverberation added to a sound and affects its recognizability by changing its spectral characteristics. Spectral content is also changed when a sound travels an indirect pathway to the ear, such as around a building. High frequency components can be absorbed by building materials and longer length, low frequency components may not travel around buildings. Sympathetic vibration can even add frequency components. This means that a familiar sound, such as M16 rifle fire, may not be clearly identified if the soldier is not familiar with the environment. Conversely, if a soldier is familiar with the environment, these changes can be informative. For example, the echoing sound of footsteps can inform a soldier that a person is inside a building, rather than outside on the sidewalk, because the concrete would not vibrate sympathetically.

Sound reaching the ears directly from the sound source contains localization cues that indicate the position of the sound source. Each time sound is reflected from a surface, it adds localization cues that indicate the position of the reflected surface. If the surface is in the same horizontal position, such as the ground or the ceiling, the localization cues will indicate the same horizontal azimuthal position. However, if the surface is in another horizontal location, the localization cues will conflict with those of the direct source.[14] The precedence effect reduces the effect of these conflicting cues, but if there is enough reverberation, or it lasts long enough, it can impair localization.

In cases where sounds must travel around a barrier, the localization cues will be consistent with the shortest pathway to the ears. As a result, the sound will appear to come from the edges of the sound barrier rather than from the original sound source.[15] Sometimes this can cause some odd effects. For example, a sound traveling through a doorway is reflected from a metal cabinet opposite the doorway. The person seated just to the side of the doorway hears the sound as coming from the cabinet, rather than the open doorway.

If sounds cause pipes or other infrastructures to vibrate, the vibration can travel along the length of the pipes. This means that vehicle traffic may be detected long before it can be recognized or identified. However, because the localization cues are emanating from the vibrating structure and not from the original sound source, it is not possible to determine the vehicle's location.

Battlefield conditions: Noise and chaos

> Noise is an important psychological weapon. The U.S. Army field manual for urban offensive operations (U.S. Department of the Army 2003) states that surprise, concentration, tempo, and audacity are especially characteristic of urban manoeuvres. Soldiers report that noise is an essential element in offensive urban operations. It can be used to surprise and startle the opposition and to convey speed and authority. For example, intense sounds (music, noise, verbal messages) played from loudspeakers mounted on low-flying helicopters or on moving vehicles may annoy and disorient the enemy as well as mask other sounds that we want to make undetectable by the enemy. Noise can also mask the onset of other sounds – an important cue for localization[16]. The sensation

[14] B. Rakerd and W. M. Hartmann, "Localization of Sound in Rooms, Iii: Onset and Duration Effects," *Journal of the Acoustical Society of America* 80 (1986): pp. 1695-706.

[15] H. Farag, J. Blauert, and O. A. Alim, "Psychoacoustic Investigations on Sound-Source Occlusion," *Journal of the Audio Engineering Society* 51 (2003): pp. 635-46.

level (SL)[17] of the target sound must be high enough[18] not only for the sound to be detected but also for the monaural and binaural localization cues to be interpretable.[19]

Multiple Sound Sources: Acoustic Distractors.

In a natural environment, there are multiple sound sources, any one of which may require attention. If two sounds occur simultaneously, it may be difficult to attend to one sound sufficiently to localize it. If two sounds occur close in time, the distractor can alter the perceived location of the sound, even if it cues that sound. The direction of this shift depends on the relative location of the sounds with respect to each other and to the listener.[20] A distractor sound located at the side will shift the estimate of the target's location away from its position, but a distractor located in the front or rear will draw the estimate nearer to its location. Multiple distractors function similarly to reverberation, adding the localization cues of each of their sources, masking and disrupting the perception of the target sound.[21]

Other Factors

---

[16] Rakerd and Hartmann.

[17] Sensation level refers to the number of decibels by which a sound exceeds a person's hearing threshold.

[18] Normal hearing listeners require a SL of about -4 to -7 dB for 50% localization accuracy within ±15° in the horizontal azimuth.

[19] Tomasz R. Letowski, Timothy Mermagen, and Kim S. Abouchacra, "Directional Detection and Localization of a Bolt Click Sound in Jungle- and Pink-Noise," In *Noise-Con 2004*. Baltimore, Maryland, 2004.

[20] K. M. Smith-Abouchacra, *Detection and Localization of a Target Signal in Multiple-Source Environments* (University Park, PA: Pennsylvania State University, 1993); J. Braasch, and K. Hartung, "Localization in the Presence of a Distracter and Reverberation in the Frontal Horizontal Plane. I. Psychoacoustical Data," *Acta Acustica United with Acustica* 88 (2002): pp. 942-55.

[21] P. M. Zurek, R. L. Reyman, and U. Balakrishnan, "Auditory Target Detection in Reverberation," *Journal of the Acoustical Society of America* 115 (2004): pp. 1609-20.

The effect of vision on auditory localization. Perhaps because sound is transient, fading away as soon as it is completed, human vision is a more reliable source of spatial location.[22] As a result, given conflicting or ambiguous visual and auditory information, the brain will sometimes misinterpret auditory cues, giving them a meaning consistent with the visual information.[23] One example of this is what is known as the "ventriloquism effect".[24] As the name implies, this phenomenon is commonly associated with the perception that the ventriloquist's "dummy" is producing the voice rather than the ventriloquist. A more general term is "visual capture", which occurs when a visual object causes an auditory stimulus to be mislocalized to the location of the visual object.[25] It seems that we are willing to trust visual location information over auditory cues. For example, it rarely concerns moviegoers that sounds accompanying events on the screen are played from loudspeakers placed to the side and the rear of the audience.

In order for visual capture to occur, the auditory and visual events must coincide in time and space.[26] Visual capture is more likely to occur when the circumstances make it a cognitively plausible interpretation of the available sensory information. For example, if only one sound source is visible, the observer is very likely to associate any sound with this source.

Under normal circumstances, when visual and auditory information are consistent, no confusion occurs. In addition, redundant sources of information usually overcome potential ambiguity. However, the soldier, dealing with ambiguity and urgency, may occasionally judge an innocuous visual object as the source of an

[22] D. R. Perrott, B. Costantino, and J. Ball, "Discrimination of Moving Events Which Accelerate or Decelerate over the Listening Interval," *Journal of the Acoustical Society of America* 93 (1993): pp. 1053-57.
[23] Y. Wada, N. Kitagawa, and K. Noguchi, "Audio-Visual Integration in Temporal Perception," *International Journal of Psychophysiology* 40 (2003): pp. 117-24.
[24] G. J. Thomas, "Experimental Study of the Influence of Vision on Sound Localization," *Journal of Experimental Psychology* 28 (1941): pp. 163-77.
[25] P. Bertelson, and M. Radeau. "Cross-Modal Bias and Perceptual Fusion with Auditory-Visual Spatial Discordance." *Perception & Psychophysics* 29 (1981): pp. 578-84.
[26] R. Bermant, and R. Welch, "Effect of Degree of Separation of Visual-Auditory Stimulus and Eye Position Upon Spatial Interaction of Vision and Audition," *Perceptual and Motor Skills* 43 (1976): pp. 487-93; Radeau and Bertelson.

alarming sound. When the soldier acts on these misperceptions, this potential adds danger to an already dangerous environment.

The visual capture effect stems from the fact that vision is superior to hearing for spatial resolution when the object is within the visual field. On the other hand, audition is superior to vision for the detection of temporal changes. For example, a single flash accompanied by two auditory beeps will be perceived as two flashes.[27] In movies, sound effects are used to draw attention to important events in a scene. In real life, sounds function in the same way. We are not able to monitor the entire visual scene constantly. Since visual objects are chiefly stationary, vigilance is simplified by attending to changes in the scene. Sound serves as a cue for some of these changes. However, when the sound environment is noisy and chaotic, we may have difficulty attending to all important events and sound becomes an important cue.

> *Moving Sound and Moving Listener.* Movement adds complexity to the situation. The effects of movement depend on who is moving - the sound source or the listener - and whether we are describing navigation through the environment or movement of the ears relative to the body.

Movements of the head while the body is otherwise stationary can help resolve binaural cues,[28] making auditory localization more accurate, especially if the sound is continuous and the sound source is stationary.[29] In addition, movement of the head may cause changes in the perceived sound spectrum due to changing monaural cues. These changes may also aid in determining the direction of incoming sound.

[27] L. Shams, Y. Kamitani, S. Thompson, and S. Shimojo, "Visual Illusion Induced by Sound," *Cognitive Brain Research* 14 (2002): pp. 147-52.

[28] F. L. Wightman, and D. J. Kistler, "Resolution of Front-Back Ambiguity in Spatial Hearing by Listener and Source Movement," *Journal of the Acoustical Society of America* 105 (1999): pp. 2841-53.
(Wightman and Kistler 1999)

[29] J. G. Fisher, and S. J. Freedman, "The Role of the Pinna in Auditory Localization," In *The Neuropsychology of Spatially Oriented Behavior*, ed. by S. J. Freedman (Homewood, IL: Dorsey Press, 1968): pp. 135-52;. A. A. Handzel, , and P. S. Krishnaprasad. "Biomimetic Sound-Source Localization." *IEEE Sensors Journal* 2 (2002): pp. 607-16; W. Noble, "Auditory Localization in the Vertical Plane: Accuracy and Constraint on Bodily Movement," *Journal of the Acoustical Society of America* 82 (1987): pp. 1631-36; S. Perrett, , and W. Noble, "The Effect of Head Rotations on Vertical Plane Sound Localization," *Journal of the Acoustical Society of America* 102 (1997): pp. 2325-32.

Humans are somewhat inaccurate at determining the absolute location of a moving sound at a particular time.[30] Consequently, estimates of the start and end-points of a moving sound are inaccurate, often biased in the direction of movement[31]. It seems that estimates of velocity are dependent on a combination of the crude estimate of the start and endpoints and the overall duration of the sound. The minimum angular displacement required to detect movement increases with velocity, suggesting that a minimum duration is required to detect and process movement. Thus, at lower angular velocities (8-16 degrees/s), movements of 3-7 degrees can be detected but larger movements are needed for detection of movement at faster velocities[32].

However, when one considers how one interacts with moving sounds and objects, this limitation seems less important. Interception of a moving sound requires some form of tracking. As long as a sound is in motion, its location is changing and precise localization is probably irrelevant. Prediction of the sounding object's future location is more useful. If movement stops, and sound continues, the listener has been alerted and can now locate the stationary signal. Errors occur when the sound ceases. Unless the listener has also been able to locate the target and see it, he must accurately estimate where the sound appeared to be when it ended and hope that it is still in this location.

If a soldier walks towards a sound object while it is sounding, he will be more accurate at determining its location than if he waited until the sound ceased[33]. As he turns towards the sound, ambiguities in the binaural cues will be removed. Then, as he

---

[30] D. R. Perrott, and A. D. Musicant, "Minimum Auditory Movement Angle - Binaural Localization of Moving Sound Sources," *Journal of the Acoustical Society of America* 62 (1977): pp. 1463-66; D. W. Grantham, "Detection and Discrimination of Simulated Motion of Auditory Targets in the Horizontal Plane," *Journal of the Acoustical Society of America* 79 (1986): pp. 1939-49.

[31] T. L. Hubbard, "Auditory Representational Momentum - Surface Form, Direction, and Velocity Effects," *American Journal of Psychology* 108 (1995): pp. 255-74; M. Nagai, K. Kazai, and A. Yagi, "Larger Forward Memory Displacement in the Direction of Gravity," *Visual Cognition* 9 (2002): pp. 28-40; S. Getzmann, , J. Lewald, and R. Guski, "Representational Momentum in Spatial Hearing," *Perception* 22 (2004): pp. 591-99.

[32] T. Z. Strybel, C. L. Manligas, and D. R. Perrott, "Minimum Audible Movement Angle as a Function of the Azimuth and Elevation of the Source," *Human Factors* 34 (1992): pp. 267-75; D. R. Perrott, B. Costantino, and J. Ball, "Discrimination of Moving Events Which Accelerate or Decelerate over the Listening Interval," *Journal of the Acoustical Society of America* 93 (1993): pp. 1053-57.

[33] D. H. Ashmead, D. L. Davis, and A. Northington, "Contribution of Listeners' Approaching Motion to Auditory Distance Perception," *Journal of Experimental Psychology - Human Perception and Performance* 21 (1995): pp. 239-56.

approaches the sound, the distance cues change. Distance cues are relative; by approaching the sound a Soldier can hear how the intensity, spectrum and offset are changing.

However, if a soldier fails to attend to the sound while it is sounding and then tries to locate it based on memory, his estimate is subject to erroneous memories of self position[34] and spatial navigation[35], decreasing the probability of an accurate guess based on auditory cues alone.

Localizability of Target Sound Sources. A sound can only be localized if it contains sufficient localization cues and these cues are audible in the environment in which it is heard. Sounds of greater than 500 ms duration, with broadband spectral content and strong onsets are easier to localize accurately than narrowband or tonal sounds with gradual onsets[36]. Overall, sounds in an urban environment are not necessarily problematic. However, sniper fire is an example of a sound that is inherently difficult to localize due to its loudness and because it often must be localized vertically as it comes from a window or from a rooftop. The fact that weapon fire is loud (157-180 dB SPL) in itself is a factor that makes it difficult to localize accurately in the vertical plane[37]. Further, the soldier is likely to hear the disturbance of air along the bullet's path and the impact of the bullet on a surface which can disrupt or bias this information. This is why microphone arrays are being developed to detect the source of sniper fire; it is very difficult to locate using human ears.

---

[34] J. R.Lackner,"The Role of Posture in Sound Localization," *Quarterly Journal of Experimental Psychology* 26 (1973): pp. 235-51; J. Lewald, G. J. Dörrscheidt, and W. H. Ehrenstein, "Sound Localization with Eccentric Head Position," *Behavioral Brain Research* 108 (2000): pp. 105-25.

[35] R. L. Klatsky, Y. Lippa, J. M. Loomis, and R. G. Golledge, "Encoding, Learning, and Spatial Updating of Multiple Object Locations Specified by 3-D Sound, Spatial Language, and Vision," *Experimental Brain Research* 149 (2003): pp. 48-61.

[36] Rakerd and Hartmann, 1985; 1986; J. Vliegen, and A. J. V. Optstal, "The Influence of Duration and Level on Human Sound Localization," *Journal of the Acoustical Society of America* 115 (2004): pp. 1705-13.

[37] R. J. Davis, and S. D. G. Stephens, "The Effect of Intensity on the Localization of Different Acoustical Stimuli in the Vertical Plane," *Journal of Sound and Vibration* 35 (1974): pp. 223-29; W. M. Hartmann, and B. Rakerd, "Auditory Spectral Discrimination and the Localization of Clicks in the Sagittal Plane," *Journal of the Acoustical Society of America* 94 (1993): pp. 2083-92; E. A. MacPherson, and J. C. Middlebrooks, "Localization of Brief Sounds: Effects of Level and Background Noise," *Journal of the Acoustical Society of America* 108 (2000): pp. 1834-49.

Familiarity

The interpretation of monaural cues requires prior knowledge of the sound. Therefore, the listener needs to be familiar with a sound in order to be able to localize a sound in elevation or distance[38]. Further, because monaural cues aid in the resolution of front-back confusions, familiarity with the sound source is also important for sound localization in the horizontal plane.

Familiarity with various sound sources and the environment itself also provide the listener with more information about which sounds to attend to in an environment. This may not improve localization accuracy specifically, but it will reduce the cognitive load by allowing one to ignore irrelevant information and, in effect, accelerate the localization process. For example, when spending the night in a new home, numerous sounds may be abnormally alerting. After a few nights in the same space, however, a person will become more accustomed to the sounds, and only sounds that are out of place will be noticed. A soldier or squad of soldiers conducting reconnaissance in an urban area is not likely to be familiar with the normal sounds or the acoustics of the environment. This will reduce situation awareness in the best of circumstances. During the course of operations, the emotional and cognitive load will reduce the interpretability of cues even further.

Summary and Recommendations

The urban environment adds acoustical and situational features that are detrimental to a soldier’s ability to detect, recognize, identify and localize relevant sound information. For the most part, both friendly and opposing forces are faced with similar limitations. However, indigenous opposing forces have a significant advantage due to familiarity with the terrain, and experience with the local sounds and the characteristics of the environment that help them better identify and locate sounds.

[38] J. W. Philbeck, and D. H. Mershon, "Knowledge About Typical Source Output Influences Perceived Auditory Distance (L)," *Journal of the Acoustical Society of America* 111 (2002): pp. 1980-83.

Unfortunately, there are limits to the sensory and cognitive capabilities of soldiers and the detrimental features of the urban battlefield cannot usually be fixed through administrative and engineering controls. However, training, practice and technology may offer some solutions that can ameliorate problems.

Auditory training with local sounds could increase familiarity with the auditory scene. Soldiers could be trained to recognize critical sound signatures such as types of weapons and vehicles. By presenting these sounds as they would sound in different urban acoustic configurations, a soldier could learn to hear the key features necessary to identify the sounding object. Further, by presenting sounds against "typical" background noises, soldiers could learn to ignore irrelevant sounds. Although "typical" environments do not exist, there are characteristic sounds of specific urban environments (e.g., the call to prayer from the minaret, sounds of vehicular traffic) that may serve as useful elements of training. It is not likely, however, that one will be able to completely adapt to a new acoustic environment without occupying it for some time. Familiarity is still important. U. S. Army doctrine wisely advises new soldiers to learn what is normal, get to know the environment and even to learn the language and culture.

Soldiers might also benefit from some simple auditory "rules of thumb". For example, movement improves sound localization if the sound is long enough (greater than 500 ms). Turning one's head towards the apparent location during the sound presentation will improve accuracy. A more advanced listener can use the presence of reverberation to answer the questions: Is this sound coming from inside or outside? Is this sound coming from that open grassy field or that wall further back?' In addition, sounds emitted in a narrow street will cause quite rapid succession of reflections unlike those from sounds emitted in an open plaza or at a market. These kinds of differences provide useful cues for orientation in urban environments.

Soldiers should also be taught about sound illusions, such as the "ventriloquist effect", so they will be aware of possible perceptual misdirection. By carefully attending to one's auditory environment and knowing the limitations of auditory localization, the soldier can develop a greater capability in identifying and localizing sounds in a given operational environment.

Technological aids for sound localization can also be developed and used. Currently, microphone arrays are used to locate sniper fire and can identify several different types of weapon signatures. Efforts are underway to make these devices smaller and more portable. There are a number of other devices that can be used to enhance hearing, such as bugs, electronic stethoscopes or a laser microphone. In addition, strategically placed sound emitting devices can minimize the information available to enemy forces by masking the sounds made by our own friendly forces.

When considering technical solutions, it quickly becomes apparent that technology can only work in a few situations and requires more planning to implement than does Soldier’s natural hearing. Thus, the soldiers’ ears and auditory training are the soldiers’ best defence.

Ashmead, D. H., D. L. Davis, and A. Northington. "Contribution of Listeners' Approaching Motion to Auditory Distance Perception." *Journal of Experimental Psychology - Human Perception and Performance* 21 (1995): 239-56.

Ashmead, D. H., D. LeRoy, and R. D. Odom. "Perception of the Relative Distances of Nearby Sound Sources." *Perception & Psychophysics* 47 (1990): 326-31.

Bermant, R., and R Welch. "Effect of Degree of Separation of Visual-Auditory Stimulus and Eye Position Upon Spatial Interaction of Vision and Audition." *Perceptual and Motor Skills* 43 (1976): 487-93.

Bertelson, P, and M. Radeau. "Cross-Modal Bias and Perceptual Fusion with Auditory-Visual Spatial Discordance." *Perception & Psychophysics* 29 (1981): 578-84.

Blauert, J. "Spatial Hearing: The Psychophysics of Human Sound Localization." Cambridge, MA: MIT Press, 1999.

Braasch, J., and K. Hartung. "Localization in the Presence of a Distracter and Reverberation in the Frontal Horizontal Plane. I. Psychoacoustical Data." *Acta Acustica United with Acustica* 88 (2002): 942-55.

Brungart, D. S., and K. R. Scott. "The Effects of Production and Presentation Level on the Auditory Distance Perception of Speech." *Journal of the Acoustical Society of America* 110 (2001): 425-40.

Chandler, D. W., and D. W. Grantham, 1992, *Journal of the Acoustical Society of America, 91*, 1624-1636.

Davis, R. J., and S. D. G. Stephens. "The Effect of Intensity on the Localization of Different Acoustical Stimuli in the Vertical Plane." *Journal of Sound and Vibration* 35 (1974): 223-29.

Divenyi, P., and J. Blauert. "On Creating a Precedent for Binaural Patterns: When Is an Echo an Echo?" In *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. Watson. New Jersey: Erlbaum, 1987.

Farag, H., J. Blauert, and O. A. Alim. "Psychoacoustic Investigations on Sound-Source Occlusion." *Journal of the Audio Engineering Society* 51 (2003): 635-46.

Fisher, J. G., and S. J. Freedman. "The Role of the Pinna in Auditory Localization." In *The Neuropsychology of Spatially Oriented Behavior*, edited by S. J. Freedman, 135-52. Homewood, IL: Dorsey Press, 1968.

Getzmann, S., J. Lewald, and R. Guski. "Representational Momentum in Spatial Hearing." *Perception* 22 (2004): 591-99.

Grantham, D. W. "Detection and Discrimination of Simulated Motion of Auditory Targets in the Horizontal Plane." *Journal of the Acoustical Society of America* 79 (1986): 1939-49.

Handzel, A. A., and P. S. Krishnaprasad. "Biomimetic Sound-Source Localization." *IEEE Sensors Journal* 2 (2002): 607-16.

Hartmann, W. M., and B. Rakerd. "Auditory Spectral Discrimination and the Localization of Clicks in the Sagittal Plane." *Journal of the Acoustical Society of America* 94 (1993): 2083-92.

Hubbard, T. L. "Auditory Representational Momentum - Surface Form, Direction, and Velocity Effects." *American Journal of Psychology* 108 (1995): 255-74.

Klatsky, R. L., Y. Lippa, J. M. Loomis, and R. G. Golledge. "Encoding, Learning, and Spatial Updating of Multiple Object Locations Specified by 3-D Sound, Spatial Language, and Vision." *Experimental Brain Research* 149 (2003): 48-61.

———. "Learning Directions of Objects Specified by Vision Spatial Audition or Auditory Spatial Language." *Learning and Memory* 9 (2002): 364-67.

Kopčo, N., and B. Shinn-Cunningham. "Effects of Cuing on Perceived Location of Auditory Sources." In *Intelligent Technologies: Theory and Applications*, edited by P. Sincak, J. Vascak, V. Kvasnicka and J. Pospichal, 201-06: IOS Press, 2002.

Lackner, J. R. "The Role of Posture in Sound Localization." *Quarterly Journal of Experimental Psychology* 26 (1973): 235-51.

Letowski, Tomasz R., Timothy Mermagen, and Kim S. Abouchacra. "Directional Detection and Localization of a Bolt Click Sound in Jungle- and Pink-Noise." In *Noise-Con 2004*. Baltimore, Maryland, 2004.

Lewald, J., G. J. Dörrscheidt, and W. H. Ehrenstein. "Sound Localization with Eccentric Head Position." *Behavioral Brain Research* 108 (2000): 105-25.

Little, A. D., D. H. Mershon, and P. H. Cox. "Spectral Content as a Cue to Perceived Auditory Distance." *Perception* 21 (1992): 405-16.

MacPherson, E. A., and J. C. Middlebrooks. "Localization of Brief Sounds: Effects of Level and Background Noise." *Journal of the Acoustical Society of America* 108 (2000): 1834-49.

McGregor, P., A. G. Horn, and M. A. Todd. "Are Familiar Sounds Ranged More Accurately?" *Perceptual and Motor Skills* 61 (1985): 1082.

Mershon, D. H., W. L. Ballenger, A. D. Little, P. L. McMurtry, and J. L. Buchanan. "Effects of Room Reflectance and Background-Noise on Perceived Auditory Distance." *Perception* 18 (1989): 403-16.

Mershon, D. H., and L. E. King. "Intensity and Reverberation as Factors in Auditory-Perception of Egocentric Distance." *Perception & Psychophysics* 18 (1975): 409-15.

Moore, B. C. J. *An Introduction to the Psychology of Hearing*. 4th ed. San Diego, CA: Academic Press, 1989.

Nagai, M., K. Kazai, and A. Yagi. "Larger Forward Memory Displacement in the Direction of Gravity." *Visual Cognition* 9 (2002): 28-40.

Nielsen, S. H. "Auditory Distance Perception in Different Rooms." *Journal of the Audio Engineering Society* 41 (1993): 755-70.

Noble, W. "Auditory Localization in the Vertical Plane: Accuracy and Constraint on Bodily Movement." *Journal of the Acoustical Society of America* 82 (1987): 1631-36.

Oldfield, S. R., and S. P. A. Parker. "Acuity of Sound Localisation: A Topography of Auditory Space. I. Normal Hearing Conditions." *Perception* 13 (1984): 581-600.

———. "Acuity of Sound Localisation: A Topography of Auditory Space. Ii. Pinna Cues Absent." *Perception* 13 (1984): 601-17.

Perrett, S., and W. Noble. "The Effect of Head Rotations on Vertical Plane Sound Localization." *Journal of the Acoustical Society of America* 102 (1997): 2325-32.

Perrott, D. R., B. Costantino, and J. Ball. "Discrimination of Moving Events Which Accelerate or Decelerate over the Listening Interval." *Journal of the Acoustical Society of America* 93 (1993): 1053-57.

Perrott, D. R., and A. D. Musicant. "Minimum Auditory Movement Angle - Binaural Localization of Moving Sound Sources." *Journal of the Acoustical Society of America* 62 (1977): 1463-66.

Philbeck, J. W., and D. H. Mershon. "Knowledge About Typical Source Output Influences Perceived Auditory Distance (L)." *Journal of the Acoustical Society of America* 111 (2002): 1980-83.

Radeau, M., and P Bertelson. "Auditory -Visual Interaction and the Timing of Inputs." *Psychological Research* 49 (1987): 17-22.

Rakerd, B., and W. M. Hartmann. "Localization of Sound in Rooms, Iii: Onset and Duration Effects." *Journal of the Acoustical Society of America* 80 (1986): 1695-706.

———. "Localization of Sound in Rooms. Ii: The Effects of a Single Reflecting Surface." *Journal of the Acoustical Society of America* 76 (1985): 524-33.

Rayleigh, Lord (J. W. Strutt, 3rd Baron of Rayleigh). "On Our Perception of Sound Direction." *Philosophical Magazine* 13 (1907): 214-32.

Shams, L, Y. Kamitani, S. Thompson, and S. Shimojo. "Visual Illusion Induced by Sound." *Cognitive Brain Research* 14 (2002): 147-52.

Smith-Abouchacra, K. M. "Detection and Localization of a Target Signal in Multiple-Source Environments." University Park, PA: Pennsylvania State University, 1993.

Steinberg, and Snow. "Physical Factors." *Bell Systems Technical Journal* 132 (1934): 245-58.

Strybel, T. Z., C. L. Manligas, and D. R. Perrott. "Minimum Audible Movement Angle as a Function of the Azimuth and Elevation of the Source." *Human Factors* 34 (1992): 267-75.

Thomas, G. J. "Experimental Study of the Influence of Vision on Sound Localization." *Journal of Experimental Psychology* 28 (1941): 163-77.

U.S. Department of the Army. "Urban Operations." Washington, D. C., 2003.

Vliegen, J., and A. J. V. Optstal. "The Influence of Duration and Level on Human Sound Localization." *Journal of the Acoustical Society of America* 115 (2004): 1705-13.

Wada, Y., N. Kitagawa, and K. Noguchi. "Audio-Visual Integration in Temporal Perception." *International Journal of Psychophysiology* 40 (2003): 117-24.

Wallach, H., E. B. Newman, and M. R. Rosenzweig. "The Precedence Effect in Sound Localization." *The American Journal of Psychology* 62 (1949): 315-36.

Wightman, F. L., and D. J. Kistler. "Resolution of Front-Back Ambiguity in Spatial Hearing by Listener and Source Movement." *Journal of the Acoustical Society of America* 105 (1999): 2841-53.

Zurek, P. M., R. L. Reyman, and U. Balakrishnan. "Auditory Target Detection in Reverberation." *Journal of the Acoustical Society of America* 115 (2004): 1609-20.