Perception of moving sources

Anna Ventura
Aalto University
Master's Program CCIS / AAT

anna.ventura@aalto.fi

Abstract

This paper investigates the role of dynamic cues in the perception of moving objects, highlighting their varying levels of significance. Findings from multiple studies emphasize the critical importance of these cues in different aspects of motion detection. Notably, these cues often work in conjunction, and their prominence varies depending on the environmental context.

Introduction

Dynamic cues are fundamental for detecting movement through sound, yet their individual effectiveness varies. This paper aims to explore the key dynamic cues relevant to perceiving moving sound sources and evaluate their respective utility. It is evident that each cue, when considered in isolation, may not effectively detect specific aspects of movement. Therefore, this paper seeks to elucidate the distinctions between these cues. The first section provides a detailed description of the various available cues and elucidates their characteristics. Additionally, it delves into objective tests and methodologies for comparing these cues, summarizing their significance and attributes.
While exploring different auditory cues, the snapshot hypothesis is repeatedly mentioned. The second section focuses on understanding this hypothesis, considering two seminal studies aimed at proving it. The first study by D. W. Grantham [1] is followed by the work of R. A. Lutfi and W. Wang [2], which not only validates the theory but also underscores sound pressure level as the primary cue for distance detection.
Furthermore, the last section describes three additional studies by C. Pörschmann, C. Störig [3], T. Kaczmarek, M. Niewiarowicz [4] and D.H. Mershon, E. King [5], which delve into the Doppler effect, sound pressure level, and interaural time difference (ITD) in greater detail, while the latter explore the usefulness of the Direct-to-reverberant cue.

Auditory cues

This section aims to elucidate the auditory cues essential for detecting sound source movement and their effectiveness. The primary cues include the Doppler effect, interaural time difference (ITD), sound pressure level (SPL), spectrum direct to reverberant energy ratio, and binaural cues. By providing an overview of these specific cues, this section seeks to enhance understanding of their significance and importance in sound perception. Considering the accuracy on detecting distance, it is important to learn how the relation between the physical distance and the perceived one is possible. It is estimated that a psychophysical power function represent the best to approximate it. [6]

$$r' = kr^a \tag{1}$$

where r' is the estimate of perceived distance, r is the actual physical source distance and k and a are fit parameters to the power function. Trough different experiments it was estimated that the average value for a is 0.4 and for k was more than one, as shown in figure 1.[6]

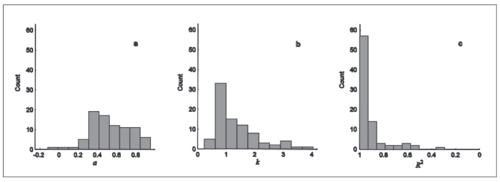
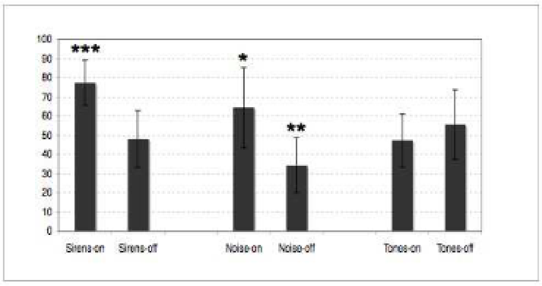


Figure 1: *Summary fit results from 84 data sets using power functions of the form $r^i = kr^a$, where r^i is an estimate of perceived distance, r is the physical source distance, and k and a are fitted parameters. Histograms of the fitted exponent, a, and constant, k , parameters are displayed in panels a and b. Panel c shows a histogram of the proportion of variance explained, R^2 , by each power function fit.*

The difficulties behind the detection of the effective distance could be multiples. One in particular could be associated with the response variability which is based on the judgement of the listener. This, even though is not entirely known, is supposedly connected to the perceptual blur typical of the auditory domain. [6]

Doppler effect

The Doppler effect stands out as a vital auditory cue, often considered the most essential. It relies on the phenomenon that alters the frequency of both the direct sound and its reflections. This phenomenon occurs naturally for listeners, tapping into our fundamental instincts. [7] Sounds influenced by the Doppler effect are perceived more effectively by humans, with approaching sounds statistically more distinguishable than those receding. As it shows in Figure 2 [8]. This heightened sensitivity may be rooted in our primitive instincts, where an approaching predator poses a greater threat than one retreating.



Black bars show the percentage of the correct identifications of directions (-on/-off) of moving sounds. Error bars represent the corresponding standard deviations. Asterisks indicate significance levels of one-sample t-tests (p<0.001(***), p<0.01(**), p<0.05(*)).

Figure 2

Interaural Time Difference, ITD

The Interaural Time Difference (ITD) is another crucial auditory cue, although research by C. J. Darwin and R. W. Hukin [9] reveals an interesting nuance. Their study suggests that subjects prioritize perceiving auditory objects in specific azimuth positions over frequency components shared by the same ITD. This implies that, depending on the context, even small differences in ITD may be necessary for accurate localization, while larger differences may not always suffice for discerning location.

Sound Pressure Level, SPL

Sound Pressure Level (SPL) is a fundamental auditory cue for discerning the location and movement of a sound source. As is commonly understood, Sound Pressure Level (SPL) is calculated using decibels, which involves comparing the absolute sound pressure to a standard reference level of sound in the air. In the section dedicated to test descriptions, insights from a study conducted by D.H. Ashmead, D. LeRoy, R.D. Odom [10] highlight the prominence of Sound pressure level as the primary auditory cue for distance detection.

Spectrum to energy Ratio

Direct-to-reverberant energy ratio is vital for gauging the distance of a sound source, especially in reverberant environments where detecting movement can be challenging. Typically, as sound moves away, its energy decreases by 6 dB when doubling the distance, while the energy of reverberant sound remains relatively constant. This discrepancy can make it difficult for listeners to accurately discern the direction and distance of sound movement[11]. However, the decreasing direct-to-reverberant energy ratio provides a crucial cue for understanding. While factors such as room reverberation and environmental conditions may introduce variability, this principle forms the basis of how this cue operates. Additionally, it has been demonstrated that the primary function of the direct-to-reverberant energy ratio cue is to furnish absolute distance information.[6]

Binaural cues

Binaural cues are pivotal for pinpointing sound sources, particularly when they are nearby. However, theoretical studies offer varying perspectives, yielding mixed information about the precise impact of binaural cues on perception and localization. Nonetheless, empirical evidence confirms the effectiveness of binaural cues for nearby sources, significantly influencing the detection of both distance and direction, particularly for lateral sources. This efficacy is attributed to the interplay of reflections and scattering phenomena between the head and torso.[6]

Snapshot Hypothesis

The snapshot hypothesis, as described by Grantham, presents a compelling perspective: "A subject, when presented with a horizontally moving target, extracts spatial samples at two discrete temporal points during the target's presentation and bases his or her judgments on the spatial difference between them" [1].
In support of this theory, R. Lufti and W. Wang [2] conducted an analysis, building upon a previous study by Rosenblum et al. (1987), which was the sole attempt at measuring cue weights in auditory motion perception.
Lufti and Wang's methodology enabled a more realistic and precise examination of this theory. They employed a different technique to simulate moving sound, encompassing discrimination of displacement, velocity, and acceleration. The primary distinction between these two studies lies in their approach to auditory cues. While R. Lufti and W. Wang compare various cues—such as Doppler, Sound Pressure Level, and interaural time difference—against one another, Rosenblum et al. (1987) opted to dynamically adjust these cues based on propagation principles. This methodology fosters a more realistic interplay among the auditory cues, enhancing the overall fidelity of the interaction.

Study

Grantham's seminal study aimed to validate the Snapshot Theory by contrasting the effects of static Minimum Audible Angle (MAA) and dynamic Minimum Audible Movement Angle (MAMA) measures. Based on the premise of the snapshot hypothesis, one would logically anticipate similar outcomes under both conditions.

To investigate this, the study standardized the duration of the sound source across settings. The experiment's results, depicted in Figure 3, unequivocally support this hypothesis. Specifically, when the sound source duration ranged between 100-150 milliseconds, the predicted parity between MAA and MAMA measures was observed. These findings affirm the validity of the snapshot theory and validate the accuracy of the predictions.

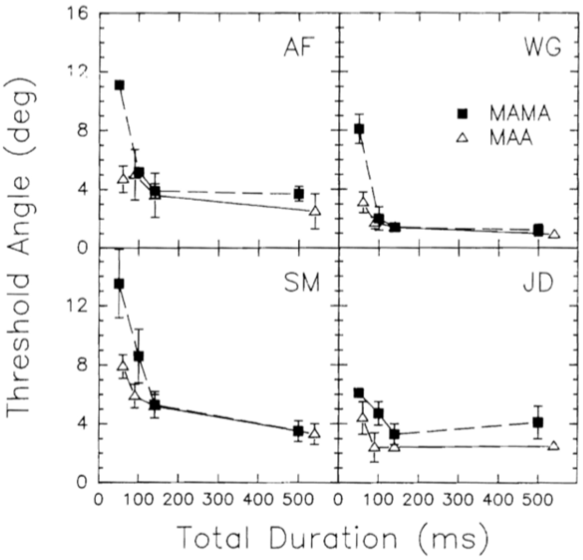


Figure 3: *Threshold in degrees as a function of the total duration of stimulus presentation. For the MAMA task duration was the presentation time of the single moving target. For the MAA task duration was the time from onset of the first marker until offset of the second marker. Medians and semi-interquartile ranges for three to six replications art shown separately for the four subjects. Where no errors are shown, the semi-interquartile ranges where smaller than the size of the datapoint.*

A more recent study to support the snapshot hypothesis is conducted by R.A. Lufti and W. Wang [2] as mentioned above. Their approach to the matter was based on a test of a moving sound, the idea was that the listener should have detect the point of minimum distance between the moving source, as shown in figure 4.

This study sought to investigate the utilization of auditory cues in the perception of velocity. Test participants were tasked with employing various auditory cues to answer four distinct questions:

1. Discrimination of a change in displacement, velocity, and acceleration
2. Effect of roving distance
3. Effect of random source spectrum
4. Discrimination at high velocity

Simulated Motion Trajectory

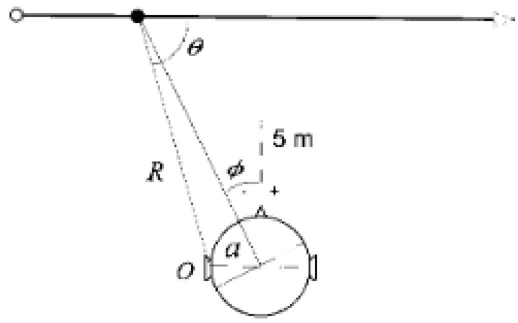


Figure 4: Simulated motion trajectory used for the standard source

The findings of this study are twofold. Firstly, they shed light on the effectiveness of different auditory cues in motion detection. Sound Pressure Level (SPL) and interaural time difference (ITD) emerge as superior cues for discerning motion changes, whereas the Doppler effect proves more adept at detecting velocity and acceleration. Secondly, these results align with those of Grantham's, previously mentioned. This consistency lends support to the Snapshot hypothesis, although with the caution that the preference for specific auditory cues yields differing results.

Test

The psychoacoustic experiments conducted by C. Pörschmann and C. Störig[3] centered around a moving car scenario in an open field, with linear motion being simulated. Two microphone setups were employed: omni-directional and a dummy head configuration.
These experiments comprised three distinct tests:

1. Investigating the influence of distance on velocity perception, while separately considering monaural- Doppler effect and SPL- and binaural- ITD- cues.

2. Examining the correspondence between perceived distance during the passing moment and the actual physical distance, also accounting for the sound source's angle.

3. Manipulating sound characteristics while keeping cues constant.

The initial study highlights the significant contribution of binaural cues, particularly in detecting movement, especially in close proximity. Results can be shown in Figure 5. The findings of the second study align closely with those of the previously mentioned test[2], providing further evidence in support of the Snapshot hypothesis.

This underscores the importance of sound pressure level in accurately determining the position of a sound source. Figure 6 shows perceived distances depending on different characteristics. Furthermore, the third study reveals that loudness serves as an additional cue for pinpointing the location of a sound source.[3] Figure 7 demonstrate how the attenuation of the stimuli influence the perception of distance.

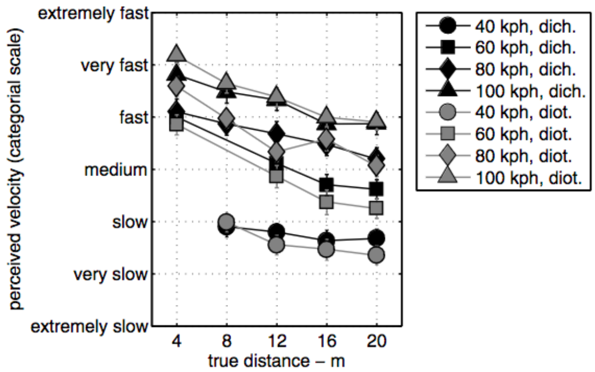
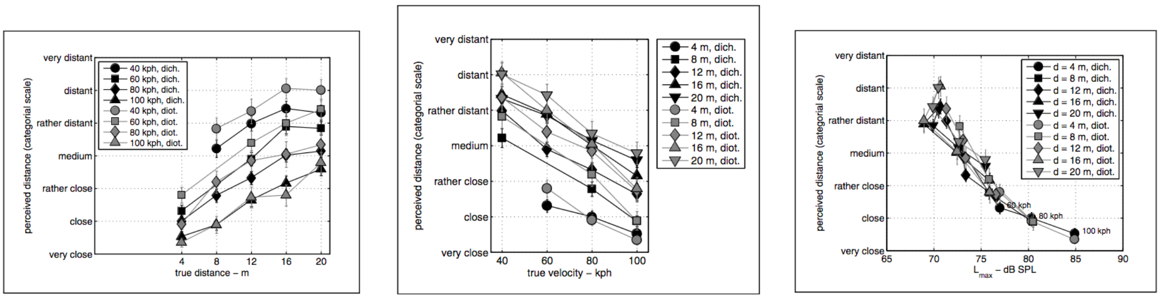


Figure 5: *Influence of the distance of the sound source on the perceived velocity.*



(a) (b) (c)

Figure 6: *Perceived distance depending on the physical distance (a); Perceived distance when passing the listener depending on the velocity of the object (b); Perceived distance when passing the listener depending on the maximum sound pressure levels at the position of the listener (c)*

In the study by T. Kaczmarek and M. Niewiarowicz[4], the simulated moving sound utilized a model based on spherical head geometry under free field conditions. A simulation of the sound of a passing car was crafted to achieve maximum realism, emphasizing careful consideration in its creation. This involved synthesizing the sound through the amalgamation of sinusoidal components, with dynamic adjustments in both level and frequency over time, ranging from 43 to 5 kHz. The spectrum utilized for this simulation is depicted in Figure 8*(a)*, while Figure 8*(b)* illustrates the experimental setup employed. Notably, the experiment catered to both individuals with normal hearing and those with impaired hearing, ensuring a comprehensive assessment of the synthesized sound's perceptual impact.

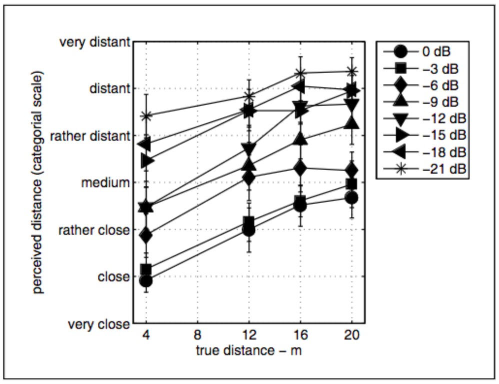
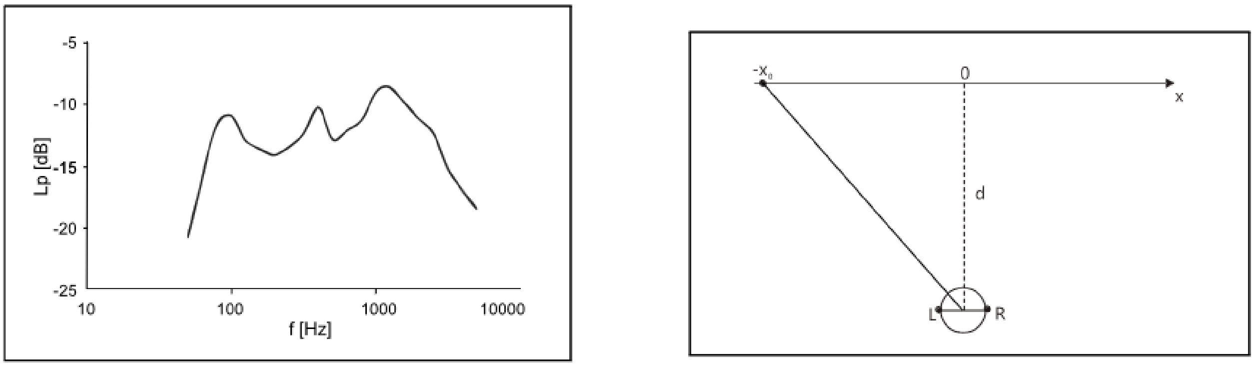


Figure 7: *Influence of stimuli attenuation on the perceived distance.*



(a) *(b)*

Figure 8: The relative spectrum of the stimulus used in the psychoacoustic experiment. *(a)* The geometry of motion used in the motion-simulations for the psychoacoustic experiments. *(b)*

Several classes of studies prioritize SPL as the primary cue when assessing distance perception. One approach involves examining the threshold of perceptible change experienced by listeners in response to variations in the proximity of the sound source. Multiple studies have demonstrated that for nearby sources, this threshold can drop to less than 6%. [10] Moreover, it is noteworthy that this percentage tends to increase for sources moving farther away from the listener, while decreasing for sources approaching closer. [10] A graphic summery of different studies is reported in figure 9.

Another set of studies were conducted where the listener would estimate the perceived sound in a matter of defined measurement. In this situation it was found that the perceived distance would increase in a lower pace than the physical distance regarding distances grater than one meter. In addition to studies that prioritize Sound pressure level as the sole auditory cue, another group of research directs its focus towards loudness. This stems from the common association between SPL and loudness, prompting an exploration into their relationship with distance. Various studies have presented contrasting findings regarding the correlation between loudness and

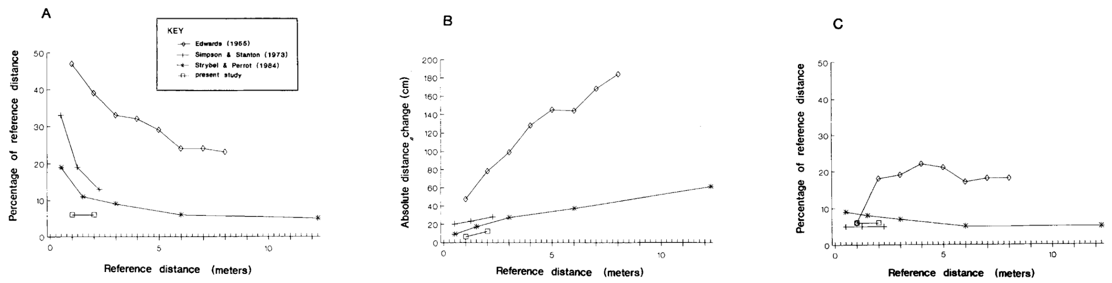


Figure 9: *Threshold values from several studies for relative distance discrimination at various reference distances. (A) Thresholds as a percentage of the reference distance, as originally reported. (B) Thresholds expressed as absolute distancechanges, also as originally reported. (C) Adjusted thresholds as a percentage of the reference distance.*

distance. Some suggest an inverse relationship [12], while others propose that they may not be intricately linked [13].In the study conducted by R. Y. Litovsky and R. K. Clifton[14], the reliance of adults on SPL as an auditory cue is questioned. The results indicate significant confusion and misinterpretation caused by variations in SPL.

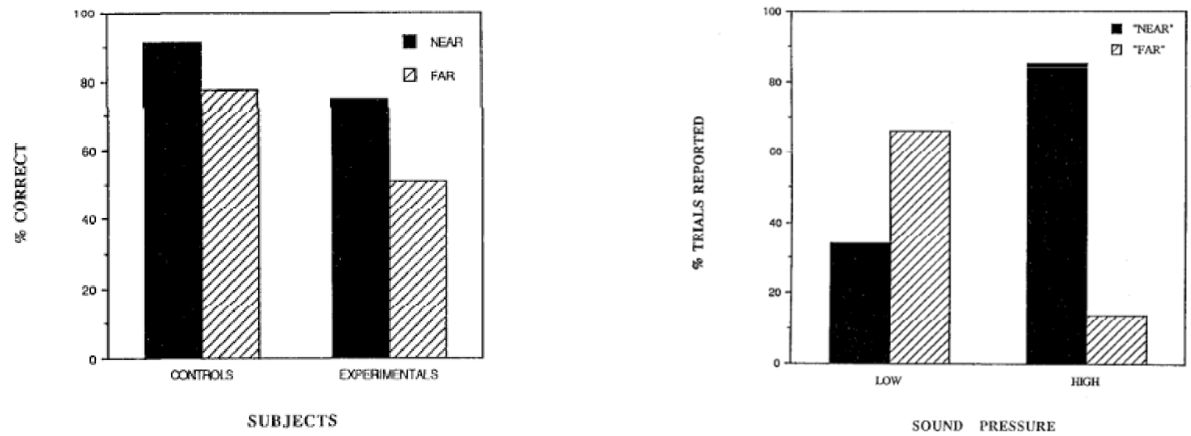


Figure 10: *Percent of trials on which adults subjects were correct in their verbal responses, plotted for control and experimental subjects, at both near and far position*

Specifically, loud sounds originating from a distance were erroneously perceived as being nearby, whereas soft sounds, despite originating from close proximity, were perceived as distant, as shown in figure 10. This underscores the pivotal role of sound pressure level in human perception and information detection. In the realm of distance detection, another significant auditory cue emerges as pivotal in listeners' ability to gauge the proximity of a sound source, especially under conditions where reverberant energy levels are low to moderate. This cue is known as the direct-to-reverberant energy ratio.

Studies conducted in reverberant environments have shown that this cue significantly enhances accuracy in distance perception compared to anechoic conditions [5]. These studies elucidate the intricate relationship between distance perception and reverberant energy. However, it's important to note that this cue may not be as reliable when it comes to relative distance judgments, but it surprisingly exhibits high accuracy in absolute distance discrimination [6].

Conclusions

After a comprehensive review of the literature, it becomes evident that certain auditory cues wield varying degrees of influence in the detection of auditory motion. Among the studies examined, there is a consensus that interaural time difference (ITD) tends to hold the greatest significance in the detection of movement of the sound source. In the study referenced as [3], particular emphasis is placed on the pivotal role of binaural effects in the estimation of velocity, contrasting with findings from [4], which downplay the relevance of binaural cues in velocity perception.
Moreover, the role of sound pressure level emerges as pivotal, particularly in discerning distance. Conversely, the Doppler effect consistently garners acclaim as the most crucial cue for accurately perceiving velocity and acceleration dynamics.In addition, it is shown that direct-to-reverberant energy ratio is essential in the detection of absolute distance, to be notice the difference between SPL which is pivotal for finest distance measures. It is possible to notice how every auditory cue can be of extreme value to detect different and particular side of the multiples needed to detect a moving source. Nonetheless, it is important to understand that take advantage of the combination is the key for a better detection of placement and movement.

References

[1] D. W. Grantham, . “Auditory motion perception: Snapshots revisited,” in Binaural and Spatial Hearing in Real and Virtual Environments, edited by R. H. Gilkey and T. R. Anderson Erlbaum, Mahwah (1997), 295–313

[2] R. A. Lutfi, W. Wang: Correlational analysis of cues for the discrimination of auditory motion. J. Acoust. Soc. Am. 106 (1999) 919–928

[3] C. Pörschmann, C. Störig: Investigations Into the Velocity and Distance, Perception of Moving Sound Sources. Acta Acustica United With Acustica 95 (2006) 696-706

[4] T. Kaczmarek, M. Niewiarowicz: Auditory Motion Perception in Normal Hearing and in Hearing Impaired People. Acta Acustica United With Acustica 99 (2013) 283 – 291

[5] D.H. Mershon, E. King: Intensity and reverberation for sensation of half distance. Percept. Mot. Skills 72 (1991) 1019-1029

[6] P. Zahorik,D. S. Brungart,A. W. Bronkhorst: Auditory Distance Perception in Humans: A Summary of Past and Present Research. Acta Acustica United With Acustic 91 (2005) 409 – 420

[7] S. Cunninghama, I. McGregorb: Evaluating use of the Doppler Effect to Enhance Auditory Alerts. Int. J. of Human–Computer Interaction 37, 11 (2021), 1074–1087

[8] M. Oechslin, M. Neukom, G. Bennett: The Doppler effect - an evolutionary critical cue for the perception of the direction of moving sound sources. ICALIP, Shanghai: IEEE (2008), 676-679

[9] C. J. Darwin, R. W. Hukin: Auditory Objects of Attention: The Role of Interaural Time Differences. Journal of Experimental Psychology: Human Perception and Performance 1999, vol. 25, No 3, 617-

[10] D.H. Ashmead, D. LeRoy, R.D. Odom: Perception of the relative distance of nearby sound sources. Percept. Psychophys. 47 (1990) 326-331

[11] P. Zahorik: Direct-to-reverberant energy ratio sensitivity. J. Acoust. Soc. Am. 112 (2002), 2110-2117

[12] S.S. Stevens, M. Guirao: Loudness, reciprocality and partiton scales. J. Acoust. Soc. Am. 34 (1962) 1466-1471.

[13] M.B Gardner: Distance estimation of 0 degrees or apparent 0 degree-oriented speech signals in anechoic space. J. Acoust. Soc. Am. 45 (1969) 47-53.

[14] R. Y. Litovsky, R. K. Clifton: Use of sound-pressure level in auditory distance discrimination by 6-month-old and adults. J. Acoust. Soc. Am. 92 (1992) 794-802