Abstract

Chipley, M. Ryan. Does auditory similarity affect the usefulness of cues to perceived relative distance? (Under the direction of Dr. Donald H. Mershon.)

This study investigated the effects of auditory similarity on cues to relative distance. Vision literature has suggested that some common visual illusions involving size contrast depend on certain physical or conceptual characteristics of the objects involved. In short, objects judged to be similar exhibit greater size contrast than objects judged to be dissimilar. This study looked at whether the same principles of similarity might also hold true for auditory distance perception. Specifically, if sounds are more similar, are they more likely to be compared to one another than are dissimilar sounds? If they are, and there are cues to relative distance available, the cues between similar sounds might be stronger. In this study, sounds were varied in spectral content, wave envelope and sound level. The stimuli consisted of 27 different sound pairs. The first sound in each pair was a broadband reference sound. The second sound was a comparison sound that varied in one of 27 different ways from the reference sound (spectral content being the same, higher, or lower; wave envelope being the same, backwards, or random; and sound level being the same, higher, or lower). The sounds were presented from a small loudspeaker located 2.5 meters in front of the listener in an acoustically "dead" room. Participants consisted of 40 students from a course in introductory psychology (20 men and 20 women) with a median age of 20. The participants used a magnitude estimation task to report perceived relative distance between the reference and comparison sounds in each sound pair. After all of the sound pairs were presented, the reference sound was presented once more by itself (either forwards or backwards). Participants gave a verbal judgment of the distance of the reference sound from themselves (i.e., an egocentric distance). The results of the study suggest that envelope and spectral content interact with sound level in determining perceived distance. Cues to relative distance were affected by "auditory similarity," but not in a systematic way. Recommendations for future research are made.

DOES AUDITORY SIMILARITY AFFECT THE USEFULNESS OF CUES TO PERCEIVED RELATIVE DISTANCE?

by
MICHAEL RYAN CHIPLEY

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

PSYCHOLOGY

Raleigh

2004

APPROVED BY:

Chair of Advisory Committee

Biography

M. Ryan Chipley was born in 1974 in Newport News, Virginia. He spent most of his childhood in northern Virginia. At 14, he and his family moved to North Carolina where he attended high school and then college. He received his undergraduate degree in psychology in 1996 from N.C. State University. After working for a few years, he returned to psychology in a graduate ergonomics/ experimental program. He and his wife, Michelle, live with their two dogs in Wake Forest, North Carolina.

Acknowledgments

This thesis would not have been possible without the help of my committee members, including my mentor and friend, Dr. Donald Mershon. Also, I want to thank my wife, Michelle, for her support, encouragement, and patience throughout the entire process.

Table of Contents

List of Tables

List of Figures

Does auditory similarity affect the usefulness of cues
to perceived relative distance?

Consider a well-established cue to relative distance—the sound level cue. It is usually described as a cue in which changes in sound level determine changes in perceived distance. However, the sounds involved almost always have the same character, be they noise, speech, tones, etc. This makes sense in that such a cue likely depends on an underlying perceptual assumption of "unity." That is, one must experience the sequence of sounds as emanating from a single sound source that moves in distance. However, suppose the second sound in a sound sequence differed from the first, not only in sound level, but in character as well. Imagine that the first sound was a noise burst presented at 50 dB, and the second sound was a man talking at 70 dB. What should we expect to happen with regard to perceived relative distance in this case? Might the dissimilarity of the sounds affect the usefulness of auditory cues to distance? Specifically, if sounds presented in sequence are not similar in character, might we expect that any differences in sound level between them will be less important, because the sounds will be judged as coming from different sources rather than from one source changing in distance?

It is often not only a single stimulus that is important in determining our perception (visual or otherwise), but also the context in which it is presented. We do not live in a simple world. We are constantly bombarded with many different kinds of stimuli in different kinds of environments. Thus, it is important to understand the interaction of multiple stimuli within the environment, in order to paint a more complete picture of spatial perception.

There are many visual phenomena that demonstrate the importance of context (e.g., color contrast, size contrast, subjective contours). A classic example of size contrast is known as the Ebbinghaus illusion. Though more-recent research has suggested that there are other factors involved than just size contrast alone (Jaeger & Guenzel, 2001), this illusion is certainly a good example of how context plays an important role in our perception of objects. Although there have been a number of versions over the years, the basic illusion consists of two equal-sized circles, each surrounded by circles of different sizes. Smaller circles surround one, and larger circles surround the other (see Fig. 1). In many studies, people reliably report that the circle surrounded by smaller circles seems larger than the circle surrounded by larger circles. The prevailing theory is that such a perception results from size contrast between the central and surrounding circles. Furthermore, the general consensus is that the magnitude of the contrast is affected by a number of factors including the similarity of the central and surrounding objects. Specifically, the more similar the central and surrounding objects are in shape and contour, the stronger the contrast between them, if there is any difference in size (Coren & Miller, 1974).

Few demonstrations exist to show the importance of context in audition. Yet, given that there are many auditory analogues to visual phenomena, it is reasonable to expect that there may be a similar effect of context in audition. As mentioned earlier, if multiple sounds are presented to a listener, changes in similar sounds may be perceived differently than changes in dissimilar sounds. Thus, judgments of relative distance may be affected.

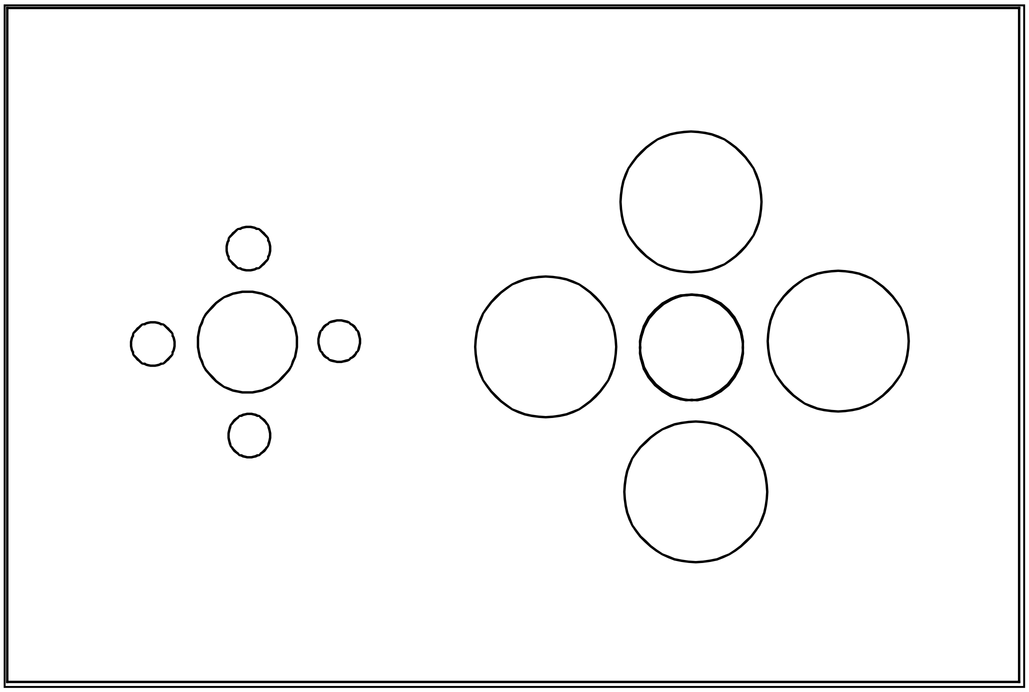

Figure 1. The Ebbinghaus Illusion. Smaller circles surround the central circle in the array on the left. Larger circles surround the central circle in the array on the right. Consequently, the left central circle (by size contrast) appears to be slightly larger than the central circle on the right.

Review

Following is a review of the research and ideas that provide the basis for the current investigation, including the Ebbinghaus illusion, the physics of sound, and cues to auditory distance.

The Ebbinghaus Illusion

The Ebbinghaus illusion has been well studied with a few particularly important investigations suggesting its underlying mechanisms. Though there has been some disagreement about exactly how and why the illusion works, a brief review of the literature is helpful in capturing the important issues involved. Coren and Miller (1974) suggested that the strength of the illusion is a function of the figural similarity of the objects that are being contrasted (i.e., the central and the surrounding figures). Figural similarity referred to different aspects of the objects, including their size and shape. Specifically, the magnitude of the illusion is significantly reduced when the surrounding objects are seen to be dissimilar from the central object.

In a subsequent study by Jaegar and Pollack (1977), the method of presentation was found to have an effect on the magnitude of the illusion. In this experiment, the center and surrounding items could be presented simultaneously or successively. When the items were presented successively, the illusion diminished. For the overestimated figure (the circle surrounded by smaller circles), successive presentation diminished the apparent size of the central circle, although not significantly. For the underestimated figure (the circle surrounded by larger circles), successive presentation significantly increased the apparent size of the central circle. Also manipulated were the relative shading and lightness of the center and surrounding figures. Lightness was found to significantly interact with both presentation and illusion type (over or underestimated). The lightness effect was greater for the underestimated figure than for the overestimated figure.

Jaeger and Grasso (1993) refuted the original size contrast and figural similarity argument for the illusion. First, figures with the largest surrounding circles generated the greatest underestimation when the surrounding circles were farthest from the central circle. If one wished to argue that it is easier to compare nearby objects than displaced ones, then this effect of separation (i.e., increasing the magnitude of the illusion) seems counter to a size-contrast explanation. Additionally, figures with like- colored central and surrounding circles did not produce larger illusions than those with unlike-colored circles. This result also does not support the argument for simple figural similarity as an explanation for the illusion.

In order to address some of the dissenting opinions in the aforementioned research and amend their earlier explanations, Coren and Enns (1993) further investigated the illusion. They found that size contrast was strongly influenced by "conceptual" similarity between central and surrounding figures. Thus, geometric similarity or visual similarity might be useful, but are only a small part of what seems to be a higher level of categorization into semantic groups. Four conditions were used to investigate conceptual similarity, including the following: *prototype* (in which the central and surrounding items were identical except in size), *same conceptual category* (in which the surrounding items were physically different from the central item, but in the same conceptual category), *near conceptual category*, and *different conceptual categories*. An example of a same conceptual category display was a drawing of a young girl's face surrounded by drawings of other young girls' faces. All fell into the same conceptual category, because all were little girls. An example of a near conceptual category was a drawing of a young girl's face surrounded by drawings of adult faces. Young girls would be placed into the conceptual category of people, but not adults. An example of a different conceptual category display was a drawing of the same young girl's face surrounded by drawings of trucks. It was found that the degree of size contrast was determined by the degree of conceptual similarity between center and surrounding figures.

A further step was taken to vary conceptual similarity independent of geometric similarity by use of some creative methodology. In this experiment, physically identical objects were used in two conditions that differed only in the orientation at which they were presented. By changing the orientation of the figures, different perceptual sets were created. Thus, geometric similarity was maintained while changing conceptual similarity. If the illusion worked according to the earlier Coren and Miller (1974) paper, then orientation should have made no difference. However, center and surrounding objects apparently coming from different categories produced a significantly smaller illusion than the identical-category (prototype) condition.

In an attempt to validate the Coren and Enns (1993) findings, Choplin and Medin (1999) investigated whether the magnitude of the Ebbinghaus illusion might provide an indirect measure of conceptual similarity. Their first experiment involved manipulation of the internal elements of complex figures. The central and surrounding objects were the same in terms of overall shape or perimeter, but differed internally (i.e., the shading and/or pattern inside each shape differed). Coren and Miller (1974) had originally manipulated only the perimeters of their objects. It was found that central and surrounding objects which were similar in perimeter yielded greater illusions than central and surrounding objects which were only internally similar. Furthermore, there was a fairly reliable monotonic trend between perimeter similarity and illusion magnitude. Another study (Jaeger & Guenzel, 2001) supported the findings regarding the importance of figural properties, rather than semantic similarity, and further suggested that the visual processes underlying size contrast include interactions of object perimeters.

Brief Review of the Physics of Sound

Loudspeakers or other objects capable of creating sound will be referred to generically as "sound sources" in this paper. The sound emitted from a point source, which emits sound energy equally in all directions, dissipates according to what is known as the inverse-square law. The inverse-square law states that sound level decreases 6 dB for every doubling of the distance from the source. This law holds for free-field situations in which the sound may radiate in space from the sound source without being impeded. Truly free fields do not exist commonly in nature, but can be approximated in a laboratory setting (e.g., in an "anechoic" chamber).

A brief discussion of sound and its characteristics is important to understanding how sound dissipates over distance. Sound is most commonly defined as the propagation of pressure waves in a medium. Most are familiar with sound traveling through the medium of air, but sound can also travel through other gases, as well as liquids and solids. Sound in air moves at an approximate velocity of 344 m/s at sea level and 20° C. However, as the elasticity of the medium increases, the speed of sound increases.

Elasticity is normally defined as the ability of a substance to recover its shape once deforming forces are removed. An example of a very elastic solid is steel. The speed of sound in steel is roughly 5000 m/s. In a less elastic substance, water, the speed of sound is roughly 1100 m/s. Changes in humidity and temperature can also cause changes in the speed of sound. Specifically, increases in temperature and humidity cause increases in the speed of sound, albeit by just a few meters per second.

Simple (pure-tone) sounds are most commonly described in terms of three different characteristics. First, a pure tone sound is a single-frequency sound whose pressure varies in a sinusoidal manner. It can be fully described in terms of its sound level, frequency, and phase (or phase relationships). Sound level, sometimes mislabeled "intensity", is based on the $\log_{10}$ of the ratio of the sound pressure of a signal to a reference pressure. It is normally measured by a sound level meter and is expressed as decibels ($dB = 20 \log_{10} P_x / P_0$). The range of decibels to which humans are sensitive is from less than 0 dB (in some cases) to more than 160 dB with no apparent upper limit except for destruction of the auditory mechanism itself.

The frequency of a tone refers to the number of times a pressure wave repeats its cycle during a standard period of time. One sinusoidal cycle includes one peak and one trough of a wave, but complex tones may have multiple peaks. The wavelength of any periodic (or repeating) wave, whether simple or complex, is the length of one cycle of the wave. Wavelength varies inversely with frequency. Higher frequency sounds have relatively shorter wavelengths than lower frequency sounds. The common way of measuring frequency is in Hertz (Hz), or cycles per second. Humans' ears are sensitive to sounds between 20 Hz and 20 KHz, with peak sensitivity in the 2 – 3 KHz range. The sounds used in this experiment will be broadband sounds or "noise" composed of multiple frequencies.

The term "phase" describes a relationship between two or more sound waves. When sound waves exist at a point simultaneously such that the peaks and troughs of each wave perfectly match-up, the waves are said to be in-phase. In this case, they reinforce each other and a higher sound level results. In any other case, where sound waves do not perfectly "match- up," waves are considered to be out-of-phase to some greater or lesser degree. In this case some cancellation will occur and will decrease over-all sound level.

Auditory Cues to Direction and Distance

Humans rely heavily on their abilities to recognize and interpret environmental information. Such information, provided by the senses, is crucial for appropriate interaction with and behavior within one's surroundings. Sound provides a wealth of environmental information. Such "localization cues" provide a listener with directional and distance information about external sound sources. The serious consideration and discussion of cues for direction is believed by many to have begun with the work of Lord Rayleigh (Rayleigh, 1907) and his description of "Duplex Theory."

Auditory Cues to Direction

Duplex theory posited that there were two primary cues for azimuth or direction--Interaural Time Difference (ITD) and Interaural Level Difference (ILD). The first of these, the interaural time difference is the delay between when a sound reaches one ear and when it reaches the other. An onset cue or an ongoing cue, each of which is useful under limited conditions, can provide such ITD information. The onset cue is useful for virtually any sound, as long as the sound to be localized has an abrupt, obvious beginning. Under conditions in which the sound to be localized does not have an obvious onset (or even if it does), an ongoing cue may prove useful. An ongoing cue can provide directional information for sounds with predominant lower frequency components (i.e., < 1500 Hz.) or other, irregular fluctuations at somewhat longer time frames. This cue has sometimes been referred to as a "phase difference" cue, because the auditory system seems to respond to the phase difference between the sound wave reaching one ear versus the other. However, the auditory system has difficulty in recognizing such phase differences for higher frequency sounds. Typical binaural delays create ambiguous patterns for sounds with short cycle durations.

The ILD cue, like the phase difference cue, is frequency dependent. It depends on the presence of *higher* frequency sounds (i.e., > 4 KHz). For such sounds, the human head is a large-enough barrier to create level differences between the ears. Such level differences are "cues" to direction. The wavelengths of sounds lower in frequency than about 4 KHz are long enough that they are able to diffract more easily around the head. In such a case, there is no discernable difference in the level of the sound reaching one ear versus the other. Thus, no level difference cue.

More modern research has contributed to the body of knowledge about auditory directional localization (for a good review, see Gilkey & Anderson, 1997). An important concept that encompasses much of what is known about how people localize sound directionally is called the HRTF (Head Related Transfer Function). It is useful in understanding the difference between the characteristics of a sound measured in a free field and the characteristics of the sound that reaches one's eardrums (Blauert, 1997).

The HRTF takes many variables into account, including directional cues, shape of the head, shape and length of the ear canals, pinna reflections, head movement, torso reflections, etc., in order to describe the transformation of a sound from air into the proximal stimulus for the ear (Brungart & Rabinowitz, 1999).

Auditory Cues to Distance

The perception of direction has traditionally been a favorite area of interest for localization researchers, but those interested in applying localization knowledge to real world problems (Loomis et al, 1999) have emphasized the importance of distance localization as well. This long-neglected area of research has flourished in recent years. A concise review of the following cues, as well as other non-acoustic factors involved in auditory distance perception, may be found in Zahorik (2002).

Sound Level Cue. In Coleman's (1963) review of cues for distance perception, he states that amplitude (loosely, "intensity") is a cue by virtue of the normal attenuation of sound with distance. Such attenuation may be explained by the inverse square law mentioned earlier. Thus, listeners will commonly perceive increases in level as being a sound that has moved closer. This relationship between changes in sound level and changes in perceived distance is well-documented (Gamble, 1909; von Békésy, 1960). While early work in this area did not make the distinction, later experiments have confirmed that such cues are "relative" (vs. absolute) in that more than a single presentation is necessary. Differences in sound level for an initial presentation of an unfamiliar sound do not reliably affect perceived distance (Mershon & King, 1975; Mershon & Bowers, 1979; Little, Mershon, & Cox, 1992).

Reverberation Cue. Reverberation can be another useful cue to distance (Bronkhorst & Houtgast, 1999). In many environments there is potential for echoes or reverberant sound. Reverberant sound is sound that has been reflected from multiple surfaces before it reaches a listener. If a sound is presented to a listener in a normal room, some of the sound will reach the listener directly (the direct sound) and some of the sound will reach the listener indirectly by way of reflection from the ceiling, floor and walls (the reverberant sound). The ratio of direct sound to reverberant sound provides information about the distance of a sound source. In general, more direct sound relative to reverberant sound indicates a shorter distance between listener and sound source (von Békésy, 1960; Blauert, 1997; McMurtry & Mershon, 1985). Studies suggest that the reverberation cue is an "absolute" cue, meaning that a single presentation of a previously unfamiliar sound is sufficient to provide a distinct percept of distance and no immediate, explicit comparison is necessary (Mershon & King, 1975; Mershon & Bowers, 1979).

Spectral Content Cue. A change in the spectral content (frequency composition) of a sound can also provide information about the distance of a sound. Such a cue to distance is based on the fact that there is a dissipation differential between high and low frequency components of a complex sound. Higher frequency sounds are more easily attenuated over distance than lower frequency sounds. Research suggests that listeners are sensitive to this differential and perceive changes in distance accordingly (Coleman, 1968). Little, Mershon, & Cox (1992) investigated whether or not experience (i.e., multiple presentations with varying content) was necessary to perceive different distances for different spectral content, and determined that indeed it was. Thus, spectral content (like sound level) should probably be considered a "relative" cue.

Purpose and Hypotheses of the Present Study

The purpose of the present experiment was to investigate a possible auditory phenomenon similar to the visual phenomenon known as the Ebbinghaus illusion. To do so, we had to make some assumptions, in order to make sense of the cross-modal comparison. First, any visual object has physical dimensions such as height, width, etc. These might be referred to as relevant descriptors of the "size" of the object. Any visual object will also have certain other characteristics that can be readily identified by sight including perimeter (or boundary), shape, type of object (person, animal, or other), color, pattern, etc. As previously indicated, many of these characteristics have been studied, in order to determine which ones contribute the most to the Ebbinghaus illusion. The general consensus seems to be that the perimeter of visual objects is the most important characteristic in determining the strength of the Ebbinghaus illusion—meaning that size contrast is strongest between objects that differ in size when the different objects have similar perimeters (or shapes). Orientation is also important in some cases, because the perimeter and size of objects might be similar (e.g., a group of equal-sized squares), but if they are not all upright (i.e., if some squares are turned so that they appear as diamonds), they will not necessarily be judged as similar. The internal components of a figure (like pattern) seem to have less of an effect on the strength of the illusion.

Similar to a visual object, an auditory "object" has certain characteristics by which it can be identified. First, a sound has a sound level which may remain constant and unchanging, or may fluctuate over time by increasing and/or decreasing. Such fluctuations in the level of a sound are commonly called its "envelope." It might also reasonably be referred to as the shape of the sound. It is possible to maintain this sound envelope while simultaneously increasing or decreasing the overall level of the sound in question. In such a case, the average sound level would change, but the fluctuation shape of the original sound would be preserved. Another characteristic of sound previously discussed is frequency. The spectral content determines the pitch and timbre (or general quality) of a sound. It is possible to vary spectral content independently of envelope. Thus, for example, two sounds could have the same overall fluctuations in level or have the same "shape," but could have different predominant frequencies – one with high frequencies and the other with low.

As mentioned before, a commonly recognized auditory cue to distance is the sound level cue. Changes in sound level cause changes in perceived distance. In order to explore the boundaries of this cue to distance, it might be useful to know how qualitatively-different sounds and qualitatively-similar sounds are judged in distance, relative to each other, when there are changes in sound level. Specifically, might we expect to find that the perceived depth separation between two similar sounds that differ in overall sound level is greater than the perceived separation between two dissimilar sounds which vary by the same level difference? If an analogy may be made between the visual Ebbinghaus literature and the current experiment, we might predict an affirmative answer to such a question. For example, if in a visual experiment, a circle and a square of different angular sizes do not appear to have depth between them, but two squares with different angular sizes do, then we might expect the same kind of thing to occur in an auditory experiment (i.e., similar sounds that differ in sound level will be more perceived to have more depth between them than dissimilar sounds under the same conditions).

What about the effect of spectral content on relative perceived distance? Changes in spectral content of a sound, due to selective attenuation of higher frequencies in air, indicate changes in distance. However, since spectral content contributes to the quality of the sound, higher frequency and lower frequency sounds may be judged as dissimilar sounds. As the sounds become less similar, relative information between them (like differences in sound level) should be less prominent. Consequently, perceived distance between such sounds would be lessened.

Whether the envelope or the spectral content of a sound is most important for determining auditory similarity remains to be seen. The present research is essentially an investigation into auditory similarity and its effect on perceived changes in distance. It is expected that a given change in sound level will produce less separation in perceived depth between perceptually dissimilar sounds than between perceptually similar sounds. Such a prediction is in line with the finding that figurally similar visual objects result in the strongest kind of Ebbinghaus illusion.

To be clear, the following are specific hypotheses regarding the outcome of the proposed experiment:

H_1 -- Decreases (increases) in level will be associated with increases (decreases) in judged depth.

H_2 -- The depth cue effect will be strongest for sound pairs which are otherwise completely matched (i.e., similar in both envelope and spectral content).

H_3 -- The depth cue effect will be stronger for sounds which match in their envelope than for those with dissimilar envelopes, spectral content being the same.

H_4 -- The depth cue effect will be stronger for sounds which match in their frequency content than for those with dissimilar frequency content, envelopes being the same.

H_5 -- The depth cue effect will be stronger for sounds which match in their envelope (but not in spectral content) than for those that match in their spectral content (but not in their envelope).

Method

General Overview

The study investigated how the similarity of sounds affected judgments of perceived relative distance. Participants made verbal judgments about the perceived distances of pairs of sounds. Each pair included an initial, unchanging standard sound and a second comparison sound. The comparison sound had either the same spectral content, wave envelope, and sound level as the standard, or it differed in one or more ways from the standard.

A total of 27 different stimulus pairs resulted from the manipulation of the three stimulus variables (spectral content, wave envelope, and sound level), each with three different values. For each presentation, the participant made a verbal judgment of the perceived distance of the comparison sound, relative to the standard sound.

Participants

The participants consisted of 40 undergraduate students (20 men and 20 women) who participated in partial fulfillment of a "research requirement" for their introductory psychology course. The median age for both men and women was 20 years. All participants were required to have normal hearing, but no audiometric testing was done.

Instead, the Hearing Screening Inventory (HSI) developed by Coren and Hakstian (1988) was used to assess overall hearing. The inventory included twelve questions related to common situations involving hearing performance and how people respond to those situations (Appendix A). In Coren and Hakstian (1988), the HSI was correlated with pure-tone hearing thresholds (PTHTs). The inventory had an internal consistency coefficient of .89 and test-retest stability coefficient of .88. It also had a correct classification rate of 92.1%. Classification of severity of hearing loss with HSI matched that of standard audiometric testing 92.1% of the time.

Environment and Apparatus

The experiment took place in a laboratory consisting of an acoustically treated test room and an adjacent control room where sounds were generated and manipulated. The test room had an entrance directly from the control room and had no windows. It had dimensions of 7.3 m X 7.3 m X 3.7 m (length X width X height). The test room contained the listener and the apparatus, and all presentations were completed in near darkness.

The walls of the test room were covered with sound-absorbent panels measuring .6 m X .6 m. Heavy pile carpeting covered the floor, in order to further reduce reverberant sound. The T_{60} measure of reverberation averaged .36 s for frequencies between .5 and 8 kHz. A room with such reverberation characteristics may be described as being acoustically "dead." The same configuration for the test room was used in (and further described in) previous studies including Mershon et al (1989).

Each listener was isolated within a small area of the test room surrounded by 2 layers of thin dark blue cotton cloth curtains. The curtains were located .6 m in front of the listener. The listener was seated in an adjustable chair, and the chair was adjusted such that the listener could rest his/her chin in a chin rest 94 cm above the floor. The chin rest was used at this height so that the average person's ears would be at a height of 102 cm (the same height as the speaker used in this study to present sounds). The experimenter instructed the listener and collected data from a position behind the curtain and just to the left of the listener. Directly behind the listener at a distance of 1.3 m, a 1.2 X 1.2 m square of 4-inch-thick Sonex® foam was mounted, in order to reduce early reflections of sound from the wall.

A wooden stand was used to support a .13 m diameter loudspeaker mounted inside a cardboard tube. The height of the center of the loudspeaker corresponded to the average expected ear height for listeners seated with their chins in the chin rest. The support structure and speaker were physically located 2.5 m in front of the listener.

Response Measures

After each stimulus pair was presented, the listener employed a magnitude-estimation task to make a verbal judgment of the perceived distance of the second sound, relative to the first. The listener was instructed that the distance of the standard was to be considered as "100." Therefore, if the listener were to perceive the second sound to be twice as far away as the standard, s/he should give a response of "200." No restrictions were placed on actual responses. Each participant made 27 such judgments (as well as additional judgments about the "egocentric" distance of the standard in feet, inches, or meters).

Using magnitude estimation minimizes the effect of different experience in the use of verbal descriptors, such as feet or inches or meters. Also, since we were primarily interested in comparing changes in perceived distance between sounds, the proportional judgments provided by the magnitude estimation task should be immediately suitable for such comparisons.

Generation, Manipulation and Measurement of Stimuli

The multiple stimuli to be used in this experiment were created by manipulating a single 4-second block of wide-band noise. The sound was manipulated in three different ways including by spectral content (same vs. high pass vs. low pass), by wave envelope (same vs. backwards vs. random), and by overall sound level (same vs. higher vs. lower). Through manipulating the aforementioned single sound, a total of 27 individual sounds were created for the experiment. These sounds were then used to create 27 pairs of sounds. Sound pairs consisted of a standard sound (to be used as a reference by the participant) and a comparison sound, each of which was 1.5 s in duration and separated from each other by 1 s of silence. Both sounds were broadband and could vary from one another in none, one, two, or all three of the dimensions mentioned earlier.

Sounds were created and manipulated using a popular sound-editing software package called "Cool Edit Pro®" (version 2.0). Each sound was stored as a WAV file on a Windows® computer (a Dell® OptiPlex GX 400 Personal Computer) located in the control room.

A sound level meter (Rion NA-61) set on fast response was used to measure the sound levels of the various sounds from the usual position of the head. The sound level of the standard sound (the sound presented at the beginning of all 27 stimulus pairs) was 58 dB (A- weighted). Each comparison sound labeled as "same" was presented at the same sound level as the standard. Comparison sounds labeled as "lower" were presented at a sound level 6 dB below the standard. Comparison sounds labeled as “higher” had a sound level 6 dB above the standard.

Octave-band measurements were taken to characterize the specific cutoffs for high and low pass filters. The results are presented in Table 1. Each of the sounds measured were 5 s samples of the wide band noise unaltered, high passed, or low passed.[1]

“Wave envelope” refers to the overall shape of the sound being presented. It represents the fluctuations of sound level within a sound over time. See Figure 2 for a graphic representation of differences in wave envelope.

Table 1: Octave band measurements for standard sound level stimuli. Measurements were taken to quantify the cutoffs for the high and low band pass filtered sounds to be used in the experiment. dB (A-weighted).

	Standard	**Low Pass**	**High Pass**
31.5 Hz	<30 dB	<30 dB	<30 dB
63 Hz	<30 dB	<30 dB	<30 dB
125 Hz	**44 dB**	**50 dB**	<30 dB
250 Hz	**48 dB**	**55 dB**	<30 dB
500 Hz	**48 dB**	**57 dB**	<30 dB
1 KHz	**52 dB**	**40 dB**	**52 dB**
2 KHz	**50 dB**	<30 dB	**50 dB**
4 KHz	**40 dB**	<30 dB	**40 dB**
8 KHz	30 dB	<30 dB	30 dB
16 KHz	<30 dB	<30 dB	<30 dB
dBA (all pass)	58 dB	58 dB	58 dB

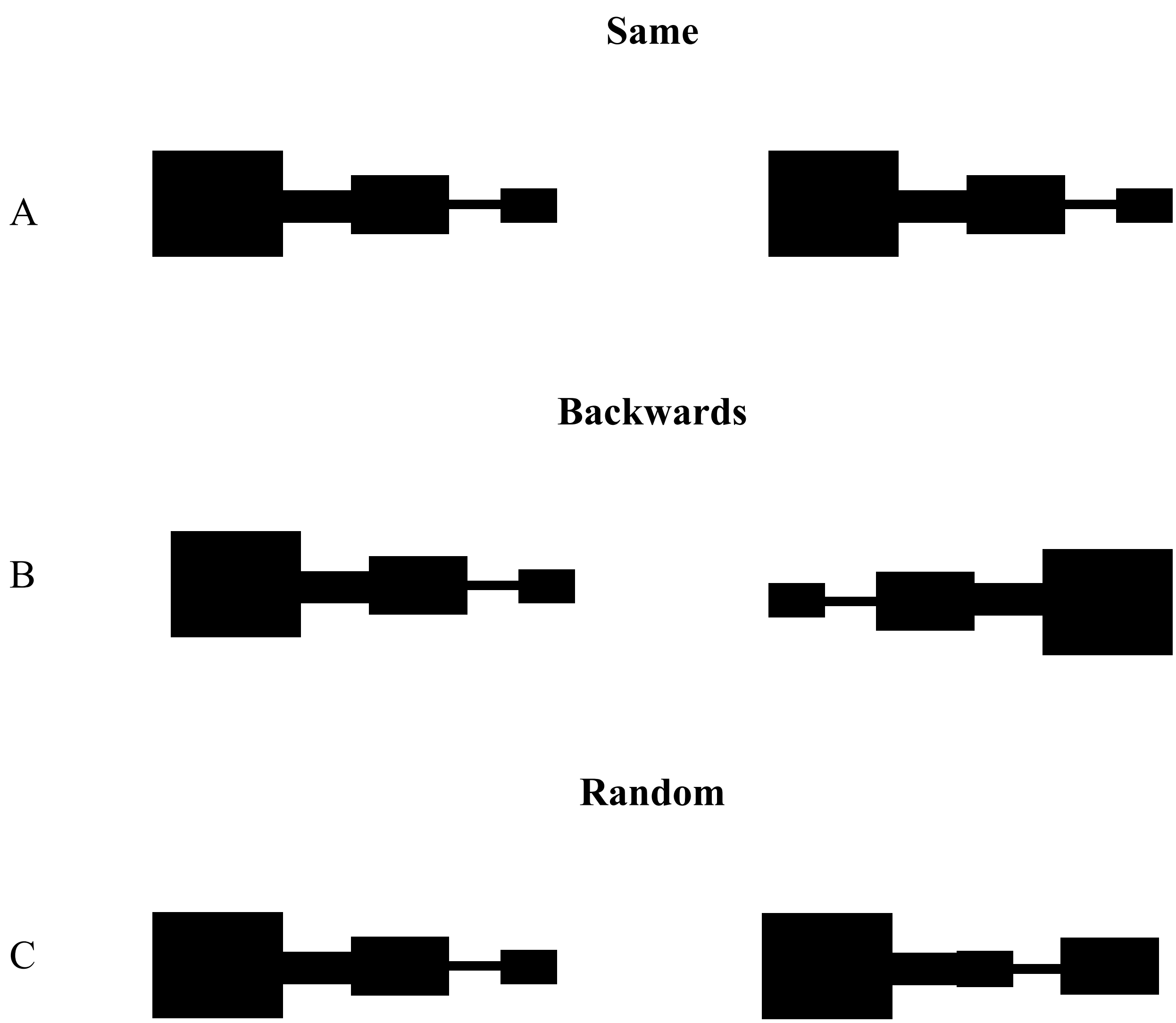

Figure 2: Wave Envelope. The above individual shapes represent sound level over time (1.5 seconds). The standard sound never changed. Thus, its shape was always the same. The comparison sound could have the same shape as the standard (A); a shape that was the standard, but with envelope reversed (B); or a shape composed of the same basic elements as the standard sound, but arranged in a random way (C). Note that all comparison sounds at a given sound level had the same RMS sound level, because they included exactly the same components.

Presentation of Stimuli

To make listeners more comfortable with the experimental task, three practice trials preceded the presentation of the first experimental stimulus pair. These practice trials consisted of comparison sounds that only varied from the reference sound in sound level. An Excel spreadsheet with embedded hyperlinks to all 27 WAV files (the stimulus pairs) was used to simplify the process of randomly ordering and presenting sounds. Each hyperlink was paired with and sorted by a random number to render a new random ordering of the stimulus pairs for each participant. Each stimulus pair was presented only once. Pairs were presented in sequence, at an inter-stimulus interval of roughly 3-5 s. This process continued until a response had been collected for each of the 27 pairs.

Design

The experimental design included 4 factors. The only between-groups factor was gender (men vs. women). The within-group factors were sound level (same, higher, lower), frequency (same, higher, lower) and wave envelope (same, reversed, scrambled).

Procedure

Each participant was greeted at the door of the laboratory and invited inside. They provided some demographic information including age, and completed the Hearing Screening Inventory described earlier. The participant read and signed the informed consent form and then read the instructions for the experiment. Any questions about the instructions were answered. The participant was escorted into the partially-darkened test room of the laboratory without a blindfold. Experimental equipment was out of view.

The listener was led to his/her position and asked to sit down in an adjustable chair. Once the listener was comfortable, with chin on chin rest, s/he was carefully blindfolded such that the ears were not blocked in any way. The experimenter moved into place behind the curtain to the left of the participant. The three practice trials were presented and any questions or concerns were addressed. The first stimulus pair was presented soon thereafter. All 27 of the stimulus pairs were presented in the same fashion, with the participant responding to each.

Before the listener was excused, s/he was presented one last time with the standard sound only, and asked to make a verbal judgment of the distance to the sound (in feet, inches, or meters). Half of the time the standard sound was presented forwards, and half of the time it was presented backwards. This was done to examine whether people were most sensitive to the first wave front of the sounds or the RMS values of the sounds when making judgments of egocentric distance. If they were using RMS, there should have been no significant difference between the forward and backward conditions. If participants used the first wave front, then the backwards condition might be judged as farther away, because its initial segment was at a lower sound level. After the experiment was completed, the participant was asked to remove the blindfold and s/he was debriefed.

Results

The first set of data collected in the experiment involved scores from Coren and Hakstian's Hearing Screening Inventory (HSI). Although everyone who participated in the experiment self- reported normal hearing, the HSI scores suggested that a total of five people (2 males and 3 females) were experiencing some hearing loss. The judged relative distance responses of the five hearing-impaired participants were not outliers or unusual in any other way. Their data was included in the final analysis as their removal made no substantive difference in the mean results.

The important variables in the experiment were sound level, spectral content, and sound envelope. All participants used magnitude estimation to judge perceived distance of a comparison sound relative to a reference sound. Their responses always reflected their perceived distance of the second sound relative to the first. As mentioned in the methods section, all participants were instructed to consider the first sound as being at a distance of "100." Thus, all responses collected were without units (see Appendix B). A couple of participants (2 men, numbered 7 and 11 in Appendix B) gave consistently extreme values for judged relative distance. These men had average judgments 3.5 standard deviations above the mean (calculated on the full data set). They were the only participants whose average distance judgments were so far from the mean. In order to correct for such extreme values, mean judgments for other non-outlier male participants were calculated. These calculated mean judgments were substituted for the outlier judgments in subsequent analyses.

Sound Level

The first hypothesis predicted that increases (decreases) in sound level should result in decreases (increases) in perceived distance. While such a result would not be surprising, it was necessary to demonstrate that the conditions and sound levels used were such that a strong depth cue was available to the participants. Without such a demonstration, investigation of the interactions between other variables (such as spectral content or envelope) and sound level would be moot. A cursory look at Figure 3 indicates that changing sound level resulted in changes in perceived distance. Indeed, analysis with a repeated-measures ANOVA indicated there was a significant main effect of sound level on perceived relative distance ($F_{2, 76} = 68.7$, $p < .0001$).

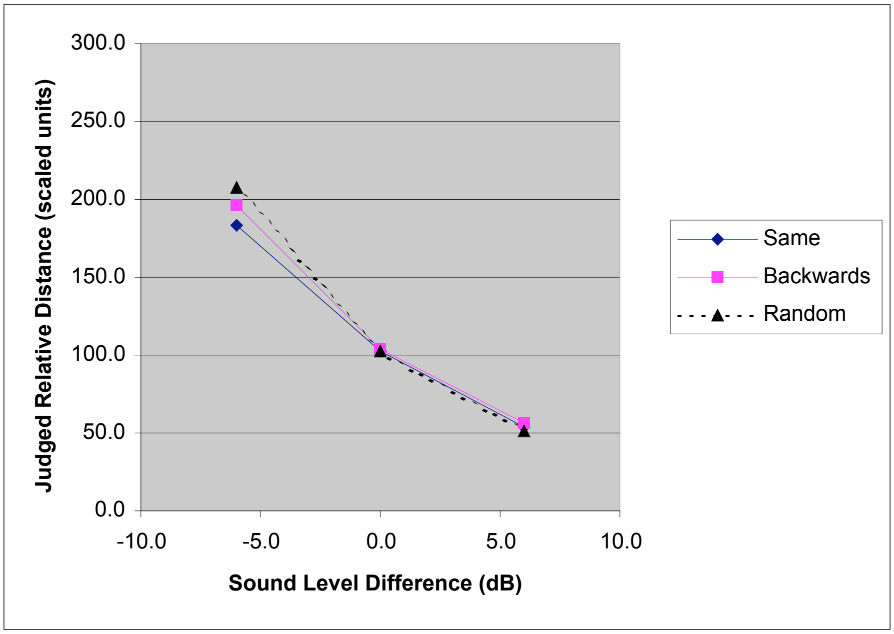

Figure 3: Judged relative distance between the reference and comparison sounds for different comparison sound wave envelopes at different sound levels.

Envelope and Sound Level

The second hypothesis of the experiment was that depth cue effects would be strongest for sound pairs which were completely matched (i.e., similar in both envelope and spectral content). Similar sounds were expected to be compared to one another more readily than dissimilar sounds. Depth information should be stronger between similar sounds. There was no significant main effect of envelope on perceived relative distance ($F_{2, 76} = .59$, $p = .56$). However, looking at Figure 3, it appears that the slopes of the plotted lines are not quite overlapping, especially at the lower sound level. This mismatch in slope represents a significant interaction between envelope and sound level ($F_{4, 152} = 10.8$, $p < .0001$). Such an interaction might suggest that auditory similarity (at least in envelope) is affecting sound level cues to distance. However, there is no support for the third hypothesis of the experiment that depth cues will be strongest for sounds which match in envelope. In fact, Figure 3 indicates that sounds dissimilar in envelope were judged as having the largest difference in perceived depth.

The fourth hypothesis involved spectral content. If sounds are similar in spectral content, then the depth cue between them should be stronger than if the sounds are dissimilar in such a way. However, Figure 4 indicates that *both* high pass *and* low pass sounds were perceived at greater distances than the reference sound when the reference and comparison sounds had the same envelope. Also, as can be seen in Figure 4, there was no difference between high and low passed sounds. One might expect such sounds to have the most perceived depth between them, because of the spectral content cue. There was indeed a significant main effect of spectral content on perceived relative distance ($F_{2, 76} = 9.1$, $p = .0003$). However, Tukey's HSD confirmed no significant difference between high and low pass sounds. It only confirmed that there were significant differences between the reference and low pass and also the reference and high pass sounds. Despite the broadband nature of the sounds, the spectral cue proved not to be as effective as it might normally be. In addition to the main effect of spectral content, a significant interaction (also evident in Figure 4) was found between spectral content and sound level ($F_{4, 152} = 3.4$, $p = .01$). Specifically, higher frequency sounds that increased in sound level by 6 dB decreased in perceived distance. They decreased more steeply than for a "same" frequency sound increasing in sound level by 6 dB. As sound level decreased for higher frequency sounds (-6 dB in Figure 4), perceived distance increased. It increased more gradually than for a "same" frequency sound with decreasing sound level. Thus, it seems that the sound level cue is less effective at indicating relationships between sounds if one sound is of sufficiently higher frequency than another. However, there is no evidence of a systematic lessening of perceived distance between sound sources that are dissimilar in frequency.

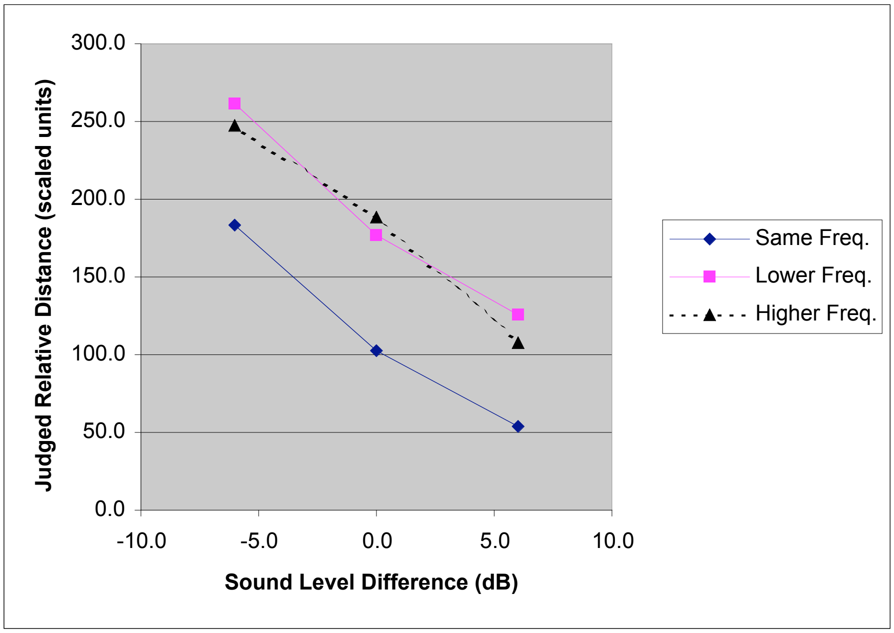

Figure 4: Judged relative distance between the reference and comparison sounds for changing comparison sound frequencies at different sound levels. Reference and comparison sound envelopes are the same.

The fifth hypothesis concerned the relative effect of envelope and spectral content on depth cues. Specifically, sounds similar in envelope (but not in spectral content) were expected to exhibit a stronger depth cue effect than for sounds that matched in spectral content (but not in envelope). As indicated by the analysis, the interaction of envelope and sound level is more effective than the interaction of spectral content and sound level. While there were significant differences between the three envelope conditions, that was not the case with the spectral content conditions. As mentioned earlier, high and low pass sounds were found to be perceptually indistinguishable from each other with regard to perceived distance. However, considering the absence of any consistent systematic variations in perceived distance resulting from manipulations in either spectral content or envelope, there is no clear difference in their influence on perceived relative depth.

Sound Level, Spectral Content, and Envelope

A significant three-way interaction was also found between sound level, spectral content, and envelope ($F_{8, 304} = 7.1$, $p < .0001$). Figures 4, 5, and 6 demonstrate the relationship between these three variables and perceived relative distance. A noteworthy finding is that, for sounds with a backwards envelope and at lower frequencies than the reference sound, increases in sound level indicated increases in distance (see Figure 5). Such a result may have contributed to the interaction depicted in Figure 3. Again, this was a demonstration that the sound level cue to distance had lost some effectiveness. However, there is no evidence to suggest something systematic. Similar sounds are not significantly farther apart with changes in sound level than dissimilar sounds with the same change in sound level.

Sex

Mean judged relative distances between men and women were not significantly different (see Table 2). However, an interaction between sex and envelope was found ($F_{2, 76} = 4.1$, $p =.02$), with more of a difference between men and women for the backward envelope condition. An interaction between sex and sound level was also present ($F_{2, 76} = 4.2$, $p =.02$), with men more responsive to changes in sound level.

Table 2: Judged relative distance (in scaled units) of sounds at lower, same, and higher sound levels than the reference sound (men vs. women).

	Men			Women		
	Lower	Same	Higher	Lower	Same	Higher
Means	95	132	185	118	171	264
Medians	70	100	150	73	118	200

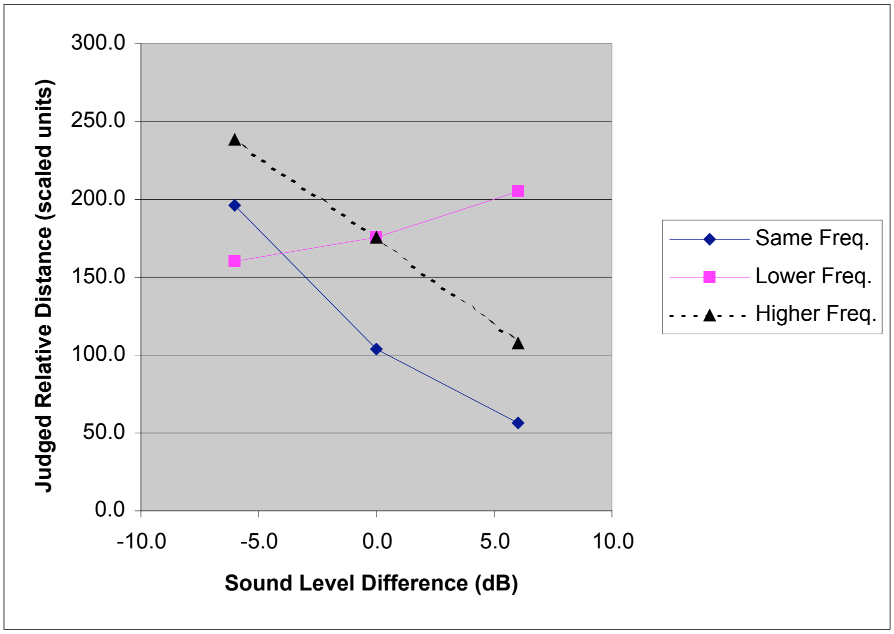

Figure 5: Judged relative distance between the reference and comparison sounds for changing comparison sound frequencies at different sound levels. The comparison sound envelope is backwards.

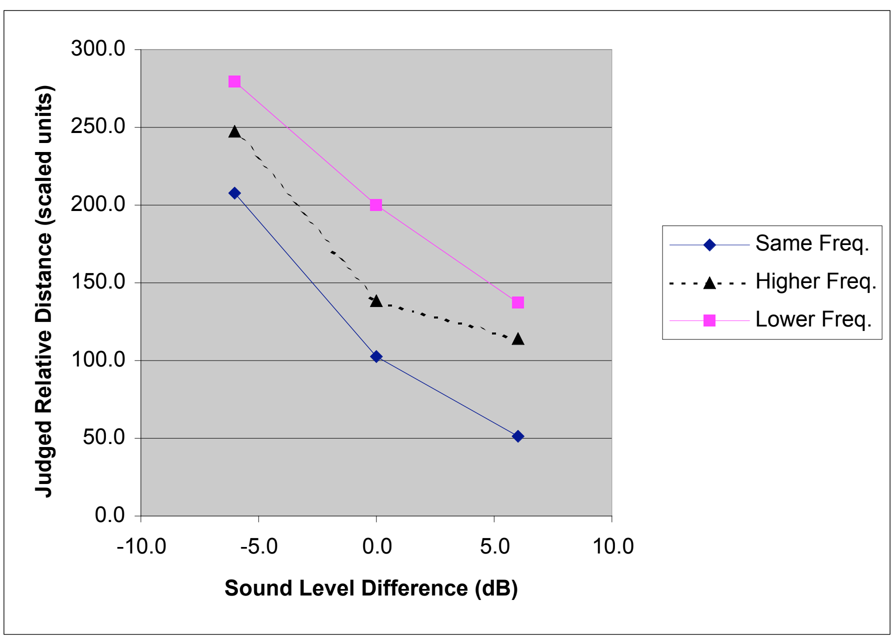

Figure 6: Judged relative distance between the reference and comparison sounds for changing comparison sound frequencies at different sound levels. The comparison sound envelope is random.

Verbal Judgments

The data collected from the experiment also included verbal judgments of the perceived distance of the reference sound. At the end of the experiment, half of the men and half of the women were presented with the reference sound itself. The others were presented with a backward reference sound. Such a measure was taken to get some idea about how listeners heard the sounds in overall egocentric distance. Even though each sound in each stimulus pair was only 1.5 s in length, the first wave front of each sound may have had special saliency in the listener's perception of distance. Since not all sounds began at the same sound level, it was important to understand whether initial sound level had any separate effect on perceptions of distance (see Table 2 for verbal reports). The results of the analysis between a forward versus a backward presentation of the reference sound using a Mann-Whitney U (U= 170, U_= 230, p< .05) suggest that such differences in presentation had no effect on judged distance. If there had been a difference, such a result would have had implications for the backward envelope stimuli used in the experiment.

Table 3: Perceived distance (in ft.) of reference sound (forward vs. backward) for men and women.

	Men		Women	
	Forward	Backward	Forward	Backward
Means	22.7 ft.	25.8 ft.	5.9 ft.	7.6 ft.
Medians	8.0	8.0	5.5	5.5
Semi-interquartile Range	4.4	5.0	0.5	2.3

Discussion

The analysis provided no conclusive evidence for a systematic effect of similarity of sounds on the effectiveness of the relative sound level cues. This is to say that sometimes dissimilarity in either frequency or envelope or both resulted in increased perceived relative distance between sounds and sometimes it resulted in reduced relative perceived distance between sounds. Sometimes such changes even resulted in no change in relative perceived distance. It is also likely that many participants perceived the backwards envelope sound as being more different from the reference sound than the random envelope sound. The backwards envelope sound started and ended at different sound levels than the reference sound. The random envelope sound began and ended at the same sound levels as the reference sound and only two segments of the random envelope sound differed from the corresponding segments in the reference sound. The difference was noticeable, but subtle.

There is no reason to suggest that the results are totally idiosyncratic. However, some results are difficult to explain. For example, in the backwards envelope sound condition, lower frequency sounds increased in distance as sound level increased. Such a result cannot be explained by any general difference between the perceived distances of backwards-envelope sounds and forwards-envelope sound. There was no such difference. However, many participants mentioned after the experiment that the low frequency sounds reminded them of (and even “sounded like”) thunder. It is very possible that familiarity cues played an unexpected role in the judgments of such sounds. Familiarity might partially explain this reversal in the sound level cue.

This was the first experiment of its kind in auditory similarity. There was thus some uncertainty about what should constitute "similarity" for sounds. Assuming that frequency and envelope can be considered as unique characteristics of a sound, there are only so many ways of varying them and there are possible confounding factors. For example, the spectral-content cue must either be overcome or at least accounted for, in order to quantify the effect of spectral-content manipulation. Also, familiarity of sounds may affect perceived distance. Although it was important in this study to keep stimuli relatively simple, in order to maintain some kind of control, this limitation (although necessary) may have been a weakness.

Generally, all of the stimuli generated for the purposes of this experiment were distinguishable, but all had the same kind of general "white noise" quality. It would be interesting to create and compare sounds that have a qualitatively more obvious difference (for example, a bird chirp, a male speaking voice, and a car engine). There are practically limitless possibilities for such stimulus combinations. Obviously, some methodological problems will have to be overcome in such an experiment. For example, it would be difficult (if not impossible) to match overall sound level for such disparate stimuli. There is also the obvious confounding influence of experience with almost any kind of complex sound. The familiar sound cue (an absolute cue to distance that depends on familiarity with a sound) would most likely exert a strong influence on perceived distance. However, that might just be the point. If we want to more closely pursue what has been discussed in the visual literature, then our definition of "similarity" will have to include more than just the most basic variables of sound. We might have to consider higher-level semantic categories of sound. That is, what kinds of sounds might be heard as belonging to the same category? Sounds that fall into "near" and "different" conceptual categories would also need to be selected. Sounds that fall into near conceptual categories could be different kinds of people speaking (e.g., a woman speaking versus a man speaking versus a child speaking). Sounds that fall into the same conceptual category could be two different women talking. Sounds that fall into different conceptual categories could be a man talking and a tiger growling. Perhaps Coren and Enns (1993) were onto something when they suggested the idea that the "conceptual" similarity of visual objects could contribute to contrast. Maybe conceptual similarity (which would depend on the use of familiar sounds) has something to do with how auditory relative cues to distance are weighted. Investigation into such a possibility is the next logical step after the present study. However, it will probably be necessary to do some pilot work in which people are given a chance to judge the similarity of multiple sounds, or include similarity judgments within the experiment itself.

Understanding auditory similarity, in the context of its effects on cues to distance, is important not only to add to our body of knowledge, but also for some practical applications. Predicting accurately where people will hear sounds in their environment is crucial for certain situations, such as virtual environments, where veridicality is key, and cockpit auditory displays, where operator survival could be at stake. The goal of either of these two examples would be to present sounds in a realistic and intuitive way. Such a presentation would likely result in better user performance. To be able to present sounds in such a way that they convey the intended information, factors that affect the usefulness of cues to distance have to be understood. Further investigation into such factors could expand and clarify the findings in the present study and contribute to the aforementioned areas of interest.

References

Békésy, G. von. (1960). Experiments in hearing. New York: McGraw-Hill.

Blauert, J. (1983). Spatial hearing. Cambridge, MA: MIT press.

Bronkhorst, A.W., & Houtgast, T. (1999). Auditory distance perception in rooms. Nature, 397, 517-520.

Brungart, D.S., & Rabinowitz, W.M. (1999). Auditory localization of nearby sources I: Head-related transfer functions. Journal of the Acoustical Society of America, 106, 1465-1479.

Choplin, J. M., & Medin, D. L. (1999). Similarity of the perimeters in the Ebbinghaus illusion. Perception & Psychophysics, 61, 3-12.

Coleman, P.D. (1968). Dual role of frequency spectrum in determination of auditory distance. Journal of the Acoustical Society of America, 44, 631-632.

Coleman, P.D. (1963). An analysis of cues to auditory depth perception in free space. Psychological Bulletin, 60, 302-315.

Coren, S., & Enns, J. T. (1993). Size contrast as function of conceptual similarity between test and inducers. Perception & Psychophysics, 54, 579-588.

Coren, S., & Miller, J. (1974). Size contrast as function of figural similarity. Perception & Psychophysics, 16, 355-357.

Gamble, E.A. (1909). Intensity as a criterion in estimating the distance of sounds. Psychological Review, 16, 416-426.

Gilkey, R. & Anderson, T. (1997). Binaural and spatial hearing in real and virtual environments. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Jaeger, T., & Grasso, K. (1993). Contour lightness and separation effects in the Ebbinghaus illusion. Perceptual and Motor Skills, 76, 255-258.

Jaeger, T., & Guenzel, N. (2001). Similarity and lightness effects in Ebbinghaus illusion created by keyboard characters. Perceptual and Motor Skills, 92, 151-156.

Jaeger, T., & Pollack, R. H. (1977). Effect of contrast level and temporal order on the Ebbinghaus circles illusion. Perception & Psychophysics, 21, 83-87.

Little, A.D., Mershon, D.H., & Cox, P.H. (1992). Spectral content as a cue to perceived auditory distance. Perception, 21, 405-416.

Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (1999) Auditory distance perception in real, virtual, and mixed environments, in Mixed reality: merging real and virtual worlds, Y. Ohta and H. Tamura, Eds., pp. 201–214. Tokyo (Japan), Ohmsha, Ltd.

McMurtry, P.L., & Mershon, D.H. (1985). Auditory distance judgments in noise, with and without hearing protection. Proceedings of the 29th Annual Meeting of the Human Factors Society, Baltimore, Maryland 811-913.

Mershon, D.H., Ballenger, W.L., Little, A.D., McMurtry, P.L., & Buchanan, J.L. (1989). Effects of room reflectance and background noise on perceived auditory distance. Perception, 18, 403-416.

Mershon, D.H., & Bowers, J.N. (1979). Absolute and relative cues for the auditory perception of egocentric distance. Perception, 8, 311-322.

Mershon, D.H., & King, L.E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. Perception & Psychophysics, 18, 409-415.

Rayleigh, L. (1907). On our perception of sound direction. Philosophical Magazine, 13, 214-232.

Zahorik, P. (2002). Auditory display of sound source distance. Proceedings of the 2002 International Conference on Auditory Displays, Kyoto, Japan, July 2-5.

Footnotes

[1] The spectral content of the sounds was evaluated with an octave-band analysis. It is possible that a narrow-band analysis would have revealed differences currently hidden within the framework of the wider analysis.

Appendix

Appendix A

Hearing Screening Inventory

Coren and Hakstian (1988)

For the first eight items, you should select the response that best describes you and your behaviors from among these response alternatives: Never (or almost never), Seldom, Occasionally, Frequently, Always (or almost always). Simply circle the letter that corresponds to the first letter of your choice. (If you normally use a hearing aid, answer as if you were not wearing it.)

1. Are you ever bothered by feelings that your hearing is poor? N S O F A
2. Is your reading or studying easily interrupted by noises in nearby rooms? N S O F A
3. Can you hear the phone ring when you are in the same room in which it is located? N S O F A
4. Can you hear the telephone ring when you are in the room next door? N S O F A
5. Do you find it difficult to make out the words in recordings of popular songs? N S O F A
6. When several people are talking in a room, do you have difficulty hearing an individual conversation? N S O F A
7. Can you hear the water boiling in a pot when you are in the kitchen? N S O F A
8. Can you follow the conversation when you are at a large dinner table? N S O F A

For the remaining four items answer using these response alternatives: Good, Average, Slightly below average, Poor, or Very poor. Again, simply circle the letter that corresponds to the first letter of your choice.

9. Overall I would judge my hearing in my right ear to be G A S P V
10. Overall I would judge my hearing in my left ear to be G A S P V
11. Overall I would judge my ability to make out speech or conversations to be G A S P V
12. Overall I would judge my ability to judge the location of things by the sound they are making alone to be G A S P V

Appendix B

Perceived relative distances for all patricipants (for each of 27 conditions)

Men

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Means	SD
Practice trials	100	100	125	100	100	100	100	75	110	100	100	100	125	100	100	90	100	100	100	110		
	300	75	75	150	125	50	75	200	150	50	50	150	175	50	50	130	50	30	50	150		
	75	150	150	50	75	100	150	50	30	200	125	75	75	200	150	70	200	150	150	40		
FsameCsameLsameNewStandard.wav	110	100	100	100	100	100	250	110	80	100	175	100	100	100	75	100	100	100	100	100	110.0	37.9
FsameCsameLlowerNewStandard.wav	275	150	125	150	125	110	800	260	120	200	600	125	200	200	125	130	50	125	150	120	207.0	179.5
FsameCsameLhigherNewStandard.wav	50	50	50	50	80	50	150	20	50	75	40	50	70	50	75	40	200	30	50	50	64.0	41.4
FsameCbackwardsLsameNewStandard.wav	125	90	100	100	95	100	400	110	110	100	275	100	140	100	75	110	100	100	100	50	124.0	77.8
FsameCrandomLsameNewStandard.wav	110	80	100	100	90	100	200	90	100	100	160	100	100	100	125	105	100	100	100	100	108.0	26.8
FsameCbackwardsLlowerNewStandard.wav	750	175	150	200	130	200	1100	300	120	125	900	150	200	200	25	170	100	120	150	140	270.3	289.3
FsameCrandomLlowerNewStandard.wav	500	150	100	200	130	100	500	150	160	200	400	150	180	200	150	140	50	150	200	120	196.5	123.9
FsameCbackwardsLhigherNewStandard.wav	75	40	25	50	50	50	300	25	50	75	200	50	70	50	75	40	50	50	50	150	76.3	66.9
FsameCrandomLhigherNewStandard.wav	75	50	50	50	75	75	200	20	10	50	200	25	70	50	90	60	200	40	30	50	73.5	58.1
FhigherCsameLsameNewStandard.wav	500	175	75	100	110	75	1200	175	120	350	500	150	250	200	125	140	150	100	150	120	238.3	258.1
FhigherCsameLlowerNewStandard.wav	900	200	75	150	175	200	1000	250	140	300	1200	150	250	400	150	180	100	115	180	150	313.3	322.5
FhigherCsameLhigherNewStandard.wav	75	30	50	50	50	50	1400	60	50	300	850	25	120	50	50	60	200	40	75	90	183.8	340.2
FhigherCbackwardsLsameNewStandard.wav	350	200	125	50	125	75	1100	80	170	300	950	175	125	200	50	170	50	75	160	120	232.5	283.2
FhigherCrandomLsameNewStandard.wav	110	115	75	50	120	125	1200	40	130	75	1300	175	110	200	150	130	50	100	75	120	222.5	354.1
FhigherCbackwardsLlowerNewStandard.wav	500	220	125	200	200	75	600	300	30	400	1900	125	240	200	100	180	25	130	300	150	300.0	404.3
FhigherCrandomLlowerNewStandard.wav	750	150	75	100	200	75	1500	500	40	400	1300	175	140	200	100	190	50	150	300	140	326.8	406.1
FhigherCbackwardsLhigherNewStandard.wav	90	60	25	50	80	75	1000	60	70	300	800	50	90	50	125	90	150	70	50	90	168.8	258.6
FhigherCrandomLhigherNewStandard.wav	100	50	25	50	80	75	1100	40	50	75	1000	50	150	50	25	95	150	30	120	80	169.8	303.7
FlowerCsameLsameNewStandard.wav	150	125	100	100	110	75	250	60	70	800	1300	125	110	100	75	210	200	80	200	100	217.0	300.8
FlowerCsameLlowerNewStandard.wav	450	175	125	150	100	75	1100	50	110	800	1800	150	200	50	175	310	50	80	350	150	322.5	438.7
FlowerCsameLhigherNewStandard.wav	250	100	75	50	100	75	1500	10	70	700	1400	125	75	50	75	240	200	25	100	70	264.5	431.9
FlowerCbackwardsLsameNewStandard.wav	500	75	100	50	125	75	1150	30	120	800	1600	150	120	50	100	230	200	50	125	110	288.0	419.3
FlowerCrandomLsameNewStandard.wav	300	100	125	50	100	75	1500	25	130	600	1600	125	110	25	75	290	200	100	140	100	288.5	450.4
FlowerCbackwardsLlowerNewStandard.wav	300	125	50	150	200	75	1200	60	150	25	1500	150	250	50	50	200	100	50	175	100	248.0	387.0
FlowerCbackwardsLhigherNewStandard.wav	900	75	150	50	70	75	1300	30	115	800	2100	125	90	50	175	300	150	70	300	160	354.3	532.4
FlowerCrandomLlowerNewStandard.wav	150	175	200	150	130	75	1500	70	200	800	1000	150	210	200	125	250	150	100	350	150	306.8	366.6
FlowerCrandomLhigherNewStandard.wav	200	90	50	50	90	75	1600	25	50	500	1500	75	80	50	100	180	200	20	80	70	254.3	455.7

Women

	1 F	2 F	3 F	4 F	5 F	6 F	7 F	8 F	9 F	10 F	11 F	12 F	13 F	14 F	15 F	16 F	17 F	18 F	19 F	20 F	Means	SD
Practice trials	100	100	100	100	100	100	125	100	110	100	100	150	98	120	150	100	100	100	100	100		
	150	50	50	150	200	50	50	150	150	50	50	200	125	90	25	180	200	50	50	200		
	50	125	150	50	45	200	200	50	50	200	200	50	50	200	125	30	50	200	200	10		
FsameCsameLsameNewStandard.wav	100	100	100	100	100	100	50	190	120	100	150	75	100	120	90	90	150	100	100	100	107.1	28.6
FsameCsameLlowerNewStandard.wav	150	135	125	125	200	200	75	150	600	200	200	200	125	300	130	200	350	200	400	250	196.7	117.7
FsameCsameLhigherNewStandard.wav	50	50	50	75	50	50	10	40	50	40	50	50	75	90	20	40	40	50	50	20	47.1	18.1
FsameCbackwardsLsameNewStandard.wav	100	100	100	100	100	100	115	100	200	100	100	110	100	150	90	90	125	100	100	100	110.4	24.5
FsameCrandomLsameNewStandard.wav	100	100	100	100	100	100	105	100	110	100	100	100	100	150	100	90	150	100	100	100	101.3	15.3
FsameCbackwardsLlowerNewStandard.wav	150	150	175	150	150	200	150	125	300	150	125	200	130	300	150	210	400	200	500	200	168.8	96.2
FsameCrandomLlowerNewStandard.wav	200	50	175	125	200	500	150	225	500	140	150	150	125	250	175	170	300	200	500	700	213.8	163.2
FsameCbackwardsLhigherNewStandard.wav	50	95	50	50	50	50	25	200	50	40	50	50	65	95	15	40	40	50	50	10	63.3	38.6
FsameCrandomLhigherNewStandard.wav	50	99	25	75	50	50	15	50	25	30	25	50	60	80	10	50	50	30	50	10	45.3	22.9
FhigherCsameLsameNewStandard.wav	150	90	150	125	125	200	75	120	500	210	50	500	140	200	150	120	250	100	300	600	191.3	149.5
FhigherCsameLlowerNewStandard.wav	200	40	250	125	300	200	200	175	650	220	125	200	150	300	200	200	600	200	300	800	223.8	186.4
FhigherCsameLhigherNewStandard.wav	100	110	75	75	75	200	25	95	200	90	50	500	85	90	50	60	150	50	50	600	132.9	145.9
FhigherCbackwardsLsameNewStandard.wav	200	110	150	125	150	200	200	105	200	200	25	500	120	250	45	190	200	200	500	500	180.4	134.3
FhigherCrandomLsameNewStandard.wav	200	90	150	125	125	200	15	115	400	140	75	300	120	200	120	170	200	100	200	350	161.3	90.7
FhigherCbackwardsLlowerNewStandard.wav	200	50	300	125	200	200	275	140	700	310	150	300	120	350	200	200	250	200	500	900	245.8	199.4
FhigherCrandomLlowerNewStandard.wav	200	50	250	125	225	200	250	200	600	210	150	500	130	300	200	200	600	200	300	800	246.7	185.7
FhigherCbackwardsLhigherNewStandard.wav	150	70	200	75	50	200	25	175	100	80	50	500	85	175	15	70	175	50	300	40	139.6	111.7
FhigherCrandomLhigherNewStandard.wav	150	100	200	75	50	50	25	110	400	200	75	400	90	90	175	50	150	50	50	600	152.9	145.1
FlowerCsameLsameNewStandard.wav	100	99	100	100	75	500	25	75	500	180	200	25	120	200	45	90	300	75	300	800	164.9	193.9
FlowerCsameLlowerNewStandard.wav	200	200	200	175	100	1000	150	150	700	180	200	900	150	200	75	160	300	200	600	600	346.3	269.1
FlowerCsameLhigherNewStandard.wav	50	90	50	75	75	25	50	90	500	200	100	20	115	90	25	60	150	50	300	300	110.4	117.3
FlowerCbackwardsLsameNewStandard.wav	150	90	75	75	75	300	50	170	600	150	150	40	150	200	80	80	400	50	100	700	160.4	178.2
FlowerCrandomLsameNewStandard.wav	50	175	175	100	50	500	125	125	600	90	200	500	125	200	30	150	400	100	300	1000	224.2	235.3
FlowerCbackwardsLlowerNewStandard.wav	50	175	50	125	70	500	10	125	300	80	200	100	200	200	115	120	350	50	200	900	148.8	198.1
FlowerCbackwardsLhigherNewStandard.wav	200	90	75	75	50	500	150	195	400	180	200	25	150	250	40	50	200	200	200	900	178.3	196.5
FlowerCrandomLlowerNewStandard.wav	200	200	200	125	150	1000	200	160	700	300	200	800	150	250	125	110	500	200	500	950	352.9	280.6
FlowerCrandomLhigherNewStandard.wav	100	90	50	75	50	500	25	75	400	180	200	10	130	90	50	70	75	75	50	1000	146.3	226.0

Appendix C

The effect of envelope, sound level, spectral content, and sex on perceived relative distance.

Source	*df*	*MS*	*F*	*p*
Envelope	2	3524	.59	.56
Envelope * Sex	2	24694	4.1	.02
Envelope * Subject (Group)	76	6006		
Sound Level	2	1280815	68.7	<.0001
Sound Level * Sex	2	77368	4.2	.02
Sound Level * Subject (Group)	76	18652		
Spectral Content	2	537761	9.1	.0003
Spectral Content * Sex	2	54998	.93	.40
Spectral * Subject (Group)	76	59072		
Spectral Content * Sound Level	4	38821	3.4	.01
Spectral Content * Sound Level * Sex	4	13934	1.2	.31
Spectral Content * Sound Level * Subject (Group)	152	11461		
Envelope * Spectral Content	4	12068	1.8	.14
Envelope * Spectral Content * Sex	4	10433	1.6	.20
Envelope * Spectral Content * Subject (Group)	152	4281		
Envelope * Sound Level	4	46287	10.8	<.0001
Envelope * Sound Level * Sex	4	1090	.26	.91
Envelope * Sound Level * Subject (Group)	152	4281		
Envelope * Spectral Content * Sound Level	8	40132	7.1	<.0001
Envelope * Spectral Content * Sound Level * Sex	8	4794	.85	.56
Envelope * Spectral Content * Sound Level * Subject (Group)	304	5617		