

ABSTRACT

CHIPLEY, MICHAEL RYAN. Does the Conceptual Similarity of Sounds Affect the Usefulness of Cues to Perceived Relative Distance? (Under the direction of Dr. Donald H. Mershon.)

This series of experiments utilized an ecological psychoacoustic approach examining the effect of conceptual similarity of sounds on cues to perceived relative distance. Traditional auditory localization research has focused on the physical aspects of sound, including things like frequency, sound level, and envelope as they relate to perception. Another popular, but entirely different, approach has been ecological in nature focusing more on environmental sounds, like speech, and testing in "real world" conditions with fewer experimental restrictions. The following series of described experiments represents a more balanced approach, incorporating aspects of traditional and ecological investigation into auditory localization.

The purpose of the experiments was ultimately to discover whether sounds more similar in conceptual category are more likely to be compared to one another than are dissimilar sounds, and if so, whether existing cues to relative distance between them might be stronger. The first of four experiments presented participants with some environmental sounds. The participants verbally identified each of the sounds. The labels for the most highly identifiable sounds were then placed on note cards. The second experiment involved participants sorting the note cards into conceptual groups/categories. A categorical analysis yielded sounds falling into same, similar, and different conceptual categories. The third experiment presented these sounds at 10 different sound levels each (from 30 – 84 dBA) to participants who reported "heard" (perceived) distances for each sound. Functions for perceived distance vs. sound level were analyzed

for each sound. Sounds with similar functions were chosen for the fourth and final experiment in which sounds from varying degrees of conceptual category were presented to listeners in pairs. Sounds in each pair also varied in sound level.

Some significant effects and interactions were found, but nothing suggesting a systematic effect of conceptual similarity on perceived relative distance in the direction expected. The most useful contribution seems to have been the perceived distance vs. sound level functions created in experiment 3. Suggestions are made for future research directions.

Does the Conceptual Similarity of Sounds
Affect the Usefulness of Cues to
Perceived Relative Distance?

by
Michael Ryan Chipley

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Psychology

Raleigh, North Carolina

2008

APPROVED BY:

James Kalat

Douglas Gillan

Christopher Mayhorn

Donald H. Mershon
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 15, 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. After working for a few years, he returned to psychology in a graduate ergonomics/ experimental program. He received his Master of Science degree in Psychology from N.C. State University in 2004. He currently works at SAS Institute, Inc. as a usability analyst.

Acknowledgements

I would like to thank my family (Dixon, Elaine, and Mark) and friends for their patience and support over the last years during my graduate studies. In addition, thanks to Marg. for her enthusiastic support. I owe a special debt of gratitude to my mentor Dr. Donald Mershon, and the rest of my committee who have seen me through this process. Also, while Dr. Larry Royster - originally a committee member, could not participate in my final defense, his contributions between my proposal and now have been valuable.

Table of Contents

List of Tables.....	vii
List of Figures.....	vii
Introduction.....	1
Review.....	3
Some Important Cues to Distance.....	3
The Ebbinghaus Illusion (A Context Effect).....	5
Details of a Previous Similarity Study.....	6
Speech as Environmental Sound.....	8
Lab versus “Real World” Sounds.....	15
Purpose and Hypotheses of the Present Study.....	17
General Overview.....	19
Experiment 1.....	20
Method.....	20
Participants.....	20
Stimuli.....	20
Procedure.....	23
Results.....	23
Discussion.....	25
Experiment 2.....	26

Method.....	26
Participants.....	26
Environment and Apparatus.....	26
Stimuli.....	26
Procedure.....	26
Design.....	27
Results.....	28
Discussion.....	33
Experiment 3.....	35
Method.....	35
Participants.....	35
Rationale.....	35
Environment and Apparatus.....	36
Stimuli.....	37
Response Measures.....	38
Design.....	38
Results.....	39
Discussion.....	41
Experiment 4.....	42
Method.....	42
Participants.....	42

Environment and Apparatus.....	42
Stimuli.....	43
Response Measures.....	44
Design.....	45
Results.....	47
Method of Analysis.....	47
Sound Level.....	48
Conceptual Category and Sound Level.....	49
Gender and Reference Sounds.....	50
Verbal Judgments.....	51
Discussion.....	53
Re-conceptualization.....	57
General Discussion.....	59
Experiment 3 Revisited.....	64
In Conclusion.....	66
References.....	68
Appendices.....	74

List of Tables

Table 1: Normalized environmental sounds from Marcell et al. (2000).....	22
Table 2: 75 normalized environmental sounds from Marcell et al. (2000).....	24
Table 3: 55 sounds yielded from the original 75 sounds in Experiment 1.....	25
Table 4: Six different presentation blocks for stimuli.....	33
Table 5: The design for Experiment 4.....	46
Table 6: Mean and median judged absolute distance (in meters) of reference sounds 1 and 2 for men and women (with semi-interquartile ranges)	51
Table 7: Intercepts and slopes for the $\Delta d'/\Delta dBA$ functions of each of the sounds from Experiment 3.....	65

List of Figures

Figure 1: Dendrogram indicating the resultant clustering of sounds.....	30
Figure 2: Dendrogram indicating the resultant clustering of sounds.....	32
Figure 3: Median perceived distances of sounds at different sounds levels.....	41
Figure 4: Mean perceived relative distances of the experimental sounds at different sound levels. All listeners combined.....	48
Figure 5: Mean perceived distances of the experimental sounds at different sounds levels with reference sound 1. All listeners combined.....	52
Figure 6: Mean perceived distances of the experimental sounds at different sounds levels with reference sound 2. All listeners combined.....	52
Figure 7: Spectrogram of "Child coughing" - the "Reference 1" sound.....	54
Figure 8: Spectrogram of "Clearing throat" - the "Reference 2" sound.....	55
Figure 9: Spectrogram of "Whistling" - the "Similar" sound.....	55
Figure 10: Spectrogram of "Sheep" - the "Different 1" sound.....	56
Figure 11: Spectrogram of "Door knocking" - the "Different 2" sound.....	56

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

Lest we become lost in the details of the following described experiments in auditory perception and forget what the point is, it's important to illustrate the purpose of such research early on. As a species we have evolved over many millennia and our sensory systems seem to have evolved in a complementary fashion (Rosenblum, 2004). There is no question that vision provides us with superior spatial resolution, with a seemingly almost isomorphic relationship between perception and "reality." Audition, on the other hand, provides us with superior temporal resolution. These differences are, in part, attributable to physiological differences between the two sensory systems. Auditory receptors are mechanical and visual receptors involve photochemical processes. Another advantage of audition is that while we can't see around corners, or behind us, or through a closed door, we can often hear- despite these restrictions. Such information is valuable should we decide to investigate things more closely with vision. In short, audition is important to our experience of the world.

Many recognize the importance of audition, and have devoted much time to research into auditory function. Traditionally, however, there has been a rift between two prominent approaches to such research. Psychoacousticians have long focused on what might be referred to as the physical aspects of sound, including things like frequency, sound level, and envelope as they relate to perception. Psychoacoustic research has a reputation for incorporating rigorous experimental controls, in order to isolate and measure particular aspects of audition. Some common laboratory restrictions in such

research include presenting sounds through headphones, or, at least, only in the horizontal plane (the plane that runs directly through the head at ear level), blindfolding listeners, and instructing listeners (or compelling them with, for example, a chin rest) to maintain a stationary head during the presentation of sounds. It is probably safe to say that the above-mentioned restrictions are artificial, in the sense that listening in the "real world" normally does not involve such restrictions. To a certain extent, that's okay. Restrictions provide us with a level of experimental control. Control promotes internal validity, and internal validity promotes sensible conclusions.

Another prominent, but entirely different, approach to auditory perception is ecological in nature. Ecological researchers have traditionally been concerned with testing in "real world" conditions. Doing research in such conditions (e.g., outside or with various naturally-occurring stimuli) allows less experimental control and, consequently, reduced internal validity. However, this sort of approach may increase a study's ecological validity. That is, we may better understand how the results apply in real environments, not just within the confines of the laboratory. This implies better generalizability, or external validity.

Most researchers would like to think that they are contributing to a body of knowledge, and that their results are somewhat generalizable across a number of conditions. However, as Mook (1983) suggests, we should be careful to consider the goals of the experiment before we determine that a study without external validity is

useless. If, for example, the goal of a study is to determine the perceptual limits of the auditory system under ideal conditions, external validity is of little importance relative to internal validity.

In the spirit of Campbell and Stanley (1963), one of the challenges of the described set of experiments was to find an optimal balance between internal and external validity. Therefore, some of the virtues of each of the aforementioned approaches were borrowed, in order to strive for what Neuhoff (2004) has called an "ecological psychoacoustics"- a notion that, until the last few years, might have been considered oxymoronic.

Review

Following is a brief review of some of the ideas and research that provide background for the current investigation, including some general information on sound cues to distance, the Ebbinghaus illusion, previous similarity studies, lessons that we can take from extensive speech research, and a consideration of the differences between lab sounds and real world sounds.

Some Important Cues to Distance

In order to understand what the goals were of the described similarity study, a brief review of sound cues to distance is helpful. One of the most important cues to distance is the sound level cue. Specifically, an increase in sound level indicates a decrease in the perceived auditory distance of a sound source. A decrease in level indicates an increase in perceived distance of the source. Thus, the sound level cue is a relative cue that depends on changes in sound level, within a (presumably) short time-

frame. Such changes in sound level indicate changes in distance (Mershon & King, 1975; Mershon & Bowers, 1979). It is relative in the sense that a listener can comment on whether a sound source is approaching or receding in distance, but has little or no information on its exact, or “absolute” distance.

The familiar sound cue is an auditory cue to distance just as is the sound level cue. However, unlike the sound level cue, it is what is known as an absolute cue – one that is available with just an isolated presentation of a particular sound. As the name “familiar” implies, it allows for the influence of past experience on the perception of the distance of a sound. If a sound is familiar, and is intense enough to be heard, a listener should only require one current presentation to make a judgment about its egocentric distance.

The familiar sound cue is worth mentioning, because in the experiment, familiar sounds were used. This cue most certainly played a role in how the sounds were heard in distance. As well, it likely interacted with effects of the sound level cue. The relative contributions of each cue are considered in the final analysis.

What about the effect of spectral content on relative perceived distance? Previous studies have indicated the importance of spectral qualities on perceived auditory distance. Such properties may influence perceived distance as both relative and absolute factors. As a relative cue, changes in the spectral content of sounds can indicate changes in distance independent of any changes in sound level (Little, Mershon, & Cox, 1992). However, it is also true that sounds with a predominance of higher-frequency content generally are perceived as closer in distance than sounds with predominantly lower-

frequency content (Butler, Levy, & Neff, 1980). On the other hand, Chipley (2004) found that both high and low-pass filtered noises were perceived as being farther away than unfiltered sounds, possibly due to simulation of familiar sounds which typically occur at remote distances. In the experiment, consideration had to be given to the fact that there were very complex spectral similarities and differences among the environmental sounds used.

The Ebbinghaus Illusion (A Context Effect)

A classic visual illusion that demonstrates the importance of context is the Ebbinghaus Illusion. 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. 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.

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, there is a popular explanation by Coren and Enns (1993). They found that the size contrast which contributed most strongly to the illusion was influenced by the “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), *similar conceptual category* (in which the surrounding items were physically different from the central item, but in a similar conceptual category), and *different conceptual categories* (in which the surrounding items were physically different from, or physically composed of geometrically identical components, but in a different conceptual category). The distinction between such categories was important for the described auditory experiments.

A previous auditory experiment (Chipley, 2004) was done to assess how some of these same context issues may be involved in hearing. In the spirit of the Ebbinghaus studies, the goal was to explore not only how we respond to a single stimulus, but also how we respond to multiple stimuli. It would be fair to say that this previous study of auditory similarity and perceived relative distance took a primarily sensory/ perceptual approach. The described experiments build on that previous study's findings by addressing some more cognitive aspects of listening to sounds.

Details of a Previous Similarity Study

Chipley (2004) investigated the effects of auditory similarity on cues to relative distance. He investigated whether sound similarity might affect auditory cues to distance. Comparisons were made between similar sounds and dissimilar sounds in perceived relative distance, in order to get a better understanding of how cues to relative distance work. In that study, sounds were varied in spectral content, wave envelope, and sound

level. All sounds were specifically designed to be unfamiliar to the listener, to avoid the confounding influence of the familiar sound cue. 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. The three independent variables were spectral content, wave envelope, and sound level.

Spectral content refers to the frequency components of a sound. Comparison sounds were either high-pass filtered (lower frequencies were removed), low-pass filtered (higher frequencies were removed), or not filtered. The labels for these different conditions were “higher,” “lower,” and “same,” respectively. “Same” refers to the fact that the comparison sounds had the same spectral content as the reference sound.

Wave envelope refers to changes in sound level within a sound. Such changes in sound level result in a difference in the temporal "shape" of the sound burst. The comparison sound had either the same shape (i.e., same changes in sound level over time) as the reference sound, or a different shape from the reference sound. If different, then the comparison sound could have an envelope that was reversed relative to the reference sound (aka “backwards”) or a wave envelope that contained all of the same level variations as the reference sound, but in a random sequence (aka "random"). The point of such variations of wave envelope was to change the overall “shape” of the sound while holding average sound levels the same, regardless of shape.

The final independent variable of sound level refers to overall sound level differences between reference and comparison sounds. The reference sound was always presented at

a sound level of 58 dB (A-Weighted). The comparison sound was either presented at the same level, at a higher level + 6 dB higher (A-Weighted), or at a level -6 dB lower (A-Weighted), referred to in the study as “same,” “higher,” and “lower,” respectively.

The sounds were presented from a small loudspeaker located 2.5 meters in front of the listener in an acoustically “dead” room (i.e., a room with low reflectivity). 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. The results of the study suggested that envelope and spectral content interact with sound level in determining perceived distance. Overall, the effectiveness of the cues to relative distance was affected by “auditory similarity,” but seemingly not in a systematic way.

Speech as Environmental Sound

As Rosenblum (2004) suggests, we can learn a lot about how to investigate auditory perception of non-speech sounds by referring to the extensive research in the area of speech. There is a lot of support for the assertion that “speech is special,” including the fact that special anatomical parts of the brain, like Broca's area and Wernicke's area, are devoted to speech processing. Theoretical functional modules like Chomsky's Language Acquisition Device also lend support to such an idea. However, the disparity between all of the answered questions about speech sounds and all of the unanswered questions about non-speech sounds may have led to the perspective that speech is perceived in a functionally different way from non-speech (Liberman & Whalen, 2000). Thus, it is to

our advantage, if we are to pursue an ecological psychoacoustics, to glean what we can from research on what is arguably the quintessential ecological sound for human beings — speech.

In the past, there has been an implicit assumption that we see *things*, but we hear *sounds* (Rosenblum, 2004). This may seem like a harmless dichotomy, but it has influenced both perceptual theory and methodology (Fowler, 1986). Gestalt ideas of whole objects have pervaded the visual literature with reference to the visual perception of objects, distance, and movement (distal properties). Yet, the focus in auditory literature, and perception textbooks like Goldstein's (2002), has been on the sensory properties of sounds like pitch, loudness, and timbre. Of course, there are also counter-examples (Mershon & King, 1975; Mershon & Bowers, 1979; Warren, 1981) that do focus on, or at least consider, distal properties. However, such examples do not dominate the literature. More importantly, Rosenblum (2004) suggests that descriptions of informational primitives for audition are often described in terms of acoustic properties, while the informational primitives for vision are described as distal object properties. For example, we recognize a musical instrument by hearing its timbre, pitch-range, and onset envelope or by *seeing* its shape, size and material. We recognize speech from hearing formants, noise bursts, and silent intervals, or *seeing* mouth shape, tongue position, and teeth. Rosenblum (2004) argues that we should consider whether the perceptual primitives for hearing might better be considered in terms of sound-structuring objects and events, than in terms of the proximal sounds themselves.

Speech research suggests that perceptual primitives correspond to articulatory gestures rather than acoustic dimensions (Fowler & Rosenblum, 1991; Liberman & Whalen, 2000). In other words, the acoustics of speech are not as important as the higher-level properties of the speech. For example, the perception of the loudness of speech, a sensory dimension, seems to be based on articulatory properties. Judgments of speech loudness correspond more closely to articulatory effort (e.g., sub glottal pressure and air velocity across the glottis) than to acoustic intensity or any other simple acoustic dimension (e.g., spectral composition) or combination of acoustic dimensions (Glave & Reitveld, 1975; Rosenblum & Fowler, 1991). For this reason, Rosenblum (2004) suggests that the perception of non-speech events may also be based on distal-event properties.

There is research that shows ability to judge event properties by virtue of sound alone, and there is research that shows that such auditory judgments correspond more closely to event properties than to acoustic primitives. One study in non-speech sound, which I will revisit later, involved the presentation by Vanderveer (1979) of 30 everyday sounds, including such sounds as hammering, footsteps, shaking keys, and tearing paper. There were three interesting results of the study. First, listeners were very good at identifying the sounds presented. Second, identification errors that were made usually made sense (e.g., footsteps and hammering were confused more often than hammering and tearing paper). Third, listeners in the study most often responded by identifying the sound-producing event, rather than by labeling the sound itself. As the listeners were instructed to simply report what they heard, such responses were made voluntarily. For

example, "shaking keys" was a much more common response than "jingling sound." The listeners seemed to resort to sensory descriptions of the sounds only when they were unable to identify the sound event.

Some more recent research has shown the ability of listeners to identify very specific characteristics of events and objects including the width of struck bars (Lakatos, McAdams, & Canssé, 1997), the length of dropped rods (Carello, Anderson, & Peck, 1997), and the hardness of struck mallets (Freed, 1990). Listeners have also shown the ability to judge whether a vessel is being filled, emptied, or maintained with different levels of fluid (Cabe & Pittenger, 2000). There are many other studies that demonstrate similar abilities. All of these studies seem to have some common attributes, including the fact that most of the sounds are familiar and the judgments that are made by listeners seem to be made with consideration of distal event properties, not simply acoustic properties. Another example study by Carello, Andersen, & Kunkler-Peck (1998) examined the parameter that could account for the "heard" length of dropped wooden rods. They found that the rod's mechanical properties, specifically the combination of each rod's elasticity and the inertial tensor were the best predictor of length judgments (accounting for 91% of the variance of heard-length judgments, regardless of rod diameter). These physical qualities were much better predictors than the acoustic amplitude, frequency, or duration of the resultant sounds. In a similar vein, there is evidence (Neuhoff, 1998) that listeners judge changes in loudness as greater for increasing intensities than for equivalent decreasing intensity, revealing what seems to be an ecological salience for approaching objects. Thus, while intensity and frequency and

their interaction will affect judgments of loudness, so may the distal event properties of *approaching* or *receding*.

The well-known *cocktail party effect* (Cherry, 1953), where a listener has difficulty attending to what one talker is saying when other talkers are speaking at the same time, has been found to vary in strength with the manipulation of speaker location information and number of speakers (Ericson, Brungart, & Simpson, 2003; Brungart & Simpson, 2007). Specifically, advance knowledge about a target-talker's location can greatly improve speech intelligibility in spatially separated, multi-talker listening conditions. Also, fewer competing speakers results in improved speech intelligibility. Unlike a typical cocktail party-type situation, in the real world, there may be multiple competing speech sources that may also be moving, more or less, continuously. This is essentially an attention problem, but a rather complex one. Which stimuli should be attended to, in order to acquire the desired information? In the case of speech, it has been found that listeners can adapt to spatial uncertainty of target-talkers in multi-talker situations, but that it requires some practice (Brungart & Simpson, 2007). Such a result may generalize to non-speech environmental stimuli. Many environmental sound stimuli have dynamically changing locations. How we attend to one or more sound targets of interest may very well correlate to how we attend to one or more sources of speech.

Furthermore, with regard specifically to Richard Warren's (1981) physical correlate theory, listeners cannot help but rely on environmental object or event primitives when making sensory judgments, because of their tacit knowledge of environmental qualities, including physical laws. For example, Warren argues that even

the most basic sensory judgments of loudness depend partly on implicit knowledge estimates of the distance of audible objects and a basic understanding of how sound changes with distance – one need not explicitly know the inverse-square law.

Similarly, other studies have indicated a listener's tendency to use non-auditory information about sound source location. For example, Mershon et al. (1981), in a series of three experiments, determined whether variations in the perceived distance of a test sound could influence its loudness in the absence of physical changes in sound-level. A “dummy” loudspeaker was used to vary the apparent source location of a short noise stimulus while the true source of this sound remained fixed (and hidden) with respect to the observer. Sound-levels from 40 to 75 dBA were presented to independent groups of observers in either anechoic or semi-reverberant acoustical environments. Reported loudness tended to increase with perceived distance, thus implying an auditory phenomenon similar to size constancy in vision – loudness constancy.

Another lesson that we can learn from speech research, according to Rosenblum (2004), is that speech often takes a higher-order form as a function of lower-level acoustic dimensions. In other words, there may be a relationship between the acoustic primitives of speech discussed earlier and higher-order form, or what Rosenblum calls gestural primitives. Although the exact relationship between the two is unclear, there is some support for the fact that gestural, and not just acoustic primitives, are important for the perception of speech. For example, in the McGurk effect, seeing the mouth of a speaker can have an important influence on what an observer may report hearing.

Rosenblum suggests that the same thing may be true with non-speech sounds. We may be just as sensitive to the events as to the acoustic primitives of the events.

Gaver (1993) has suggested that we must examine sound source mechanics, in order to attain a complete understanding of true acoustic information. In a seminal paper in this area, he considered different types of events with an emphasis on the mechanical components likely to have an audible consequence: impacts, scrapes, and drips. He found that the physical differences in vibratory behavior between struck-wooden and struck-metal bars involve different damping functions, which have different acoustic outcomes. The decay of a metal bar's signal is slower overall, with higher frequency components lasting longer than lower frequency components. It was just the opposite with the wooden bar, which had a faster overall decay time, with lower frequencies lingering longer than higher ones. A listener might very well be sensitive to such acoustic differences and distinguish between the two separate materials based on such acoustic differences.

Other evidence suggests that, similar to speech, time-varying components of a non-speech signal are salient. Warren and Verbrugge (1983) studied listeners' ability to recognize the difference between a bouncing versus a breaking jar. One might think that the most obvious difference would be the initial crashing sound of the breaking jar. However, when the initial impact noise was removed, listeners were still able to tell the difference between the two. Furthermore, when the spectral structure from one event was applied to the temporal patterning of the other event, it was the temporal patterning that mainly determined the identification of the event.

To summarize the previous literature, we should consider non-speech sounds as events that are based on higher-order, time-varying, production-based parameters. Much as we see objects and not separate sensory qualities, we hear events and not sounds *per se*. The acoustic information that is available to us from sound events is interpretable, but generally our perception of said events cannot be explained exclusively by reference to any one, or sometimes more than one, acoustic feature such as sound amplitude, spectral content, or wave envelope.

Lab versus "Real World" Sounds

Whether people are more or less accurate in their judgments of the distance of lab created sounds (like white noise) versus familiar "real-world" sounds is an important question. One study has shown errors occur more often with "real world" sounds (Barneclutt & Pfeffer, 1998). This is not surprising, since there are countless and unpredictable auditory events in the real world – some of which may have undue influence on both auditory directional and distance perception. However, this was not an issue addressed experiments. We were interested instead in comparing "real-world" (or environmental) sounds to each other and not to laboratory sounds. In doing so, the hope was to get a better understanding of how auditory distance cues work together with our cognitive faculties to determine the distances of different sounds.

The identification of environmental sounds and judgments of their similarity are areas that have been investigated (Ballas, 1993; Bonebright, 2001; Gygi, 1997; Gygi et al., 2004; Marcell et al., 2000). It is interesting to consider what characteristics of environmental (i.e., non-speech) sounds a listener is attuned to when identifying such a

sound? This is really the type of question that a psychoacoustician would be eager to address. Gygi et. al (2004) found that the answer to this question depends on a number of factors. Specifically, most sounds can be easily identified by information in the 600 Hz - 4800 Hz range. This is not completely surprising, given the fact that humans are most sensitive to sounds within the range of about 1 - 4 KHz. Gygi also found that some sounds could be identified without any frequency information at all, just by relying on the temporal structure of the sounds. Such temporal features that may be important are the number of peaks, amount and distribution of silence, and other features.

There are a number of ideas about how to compare environmental sounds psychoacoustically, but the methods are not standardized and can be very complex. Such quantification of complex sounds is outside the scope of this investigation. In fact, many studies have shown that no single dimension can predict certain judgments about sound events or objects. In such studies, it is likely that some higher-order combination of acoustic dimensions is most important. By "higher-order," I am not referring to a simple interaction between dimensions, but rather a complex interaction requiring sophisticated statistical modeling. The purpose of the described investigation was not to define those higher-order combinations in terms of the acoustical minutia of sounds and the interactions of those complicated quantifications. Much as if we were doing visual-similarity research, we are not interested in categorizing the stimulation of retinal ganglion cells (or that of cells in different parts of the occipital lobe) by the lengths, widths, colors, corners, reflectance, etc. of the lines that compose the visual object or objects at which we are looking. To avoid the complex spectral analyses and

comparisons necessary to quantify each of the environmental sounds used, the described set of studies simply strove to manipulate sound level and semantic similarity.

Appropriate starting sound levels for each sound will be determined with experiment 2.

Purpose and Hypotheses of the Present Study

The purpose of the described experiments was to further investigate some of the findings in Chipley (2004). In that study, some assumptions were made about "auditory objects" and what constituted similarity based on findings in the literature on visual similarity. The acoustic characteristics deemed to be suitable as "features" of a sound were envelope and spectral content. The literature reviewed earlier suggests that such a "time-varying" aspect of sound might be salient. Spectral content refers to the frequency content of a sound. Given the paucity of other auditory-similarity research, these assumptions were a logical place to start. However, given the findings in Chipley (2004), other characteristics of sounds, as well as the use of more-complex sounds, need to be considered, especially if the results are going to be ecologically (and externally) valid.

In the spirit of the study by Coren and Enns (1993), the experiments considered sounds that fell into different conceptual-similarity categories. The same categories as used by Coren and Enns were utilized: *prototype*, *same*, *similar*, and *different*. The *prototype* case involved two sounds that were exactly the same in every way, except for differences in overall sound level. Sounds that fell into the *same* category were sounds that were judged as conceptually part of the same group of objects (e.g., two bouncing objects, like a basketball and a tennis ball). Sounds that fell into *similar* conceptual categories were sounds that were judged as part of similar conceptual groups (e.g., a

faucet turning and water running). Finally, sounds that fell into different conceptual categories were sounds that were judged to be in completely different conceptual categories (e.g., glass breaking and footsteps). The first of four experiments was conducted, in order to assess whether people consistently identified the sounds to be used in the experiments that followed. In exp. 2, participants were asked to group the sounds into categories. Participants received a deck of cards with each card containing a description of a sound, like a car horn or a dog bark (or any other sound previously determined as sufficiently identifiable in the first experiment) and asked to indicate their judgments about the similarity of those sounds by sorting the cards into different piles. Hierarchical cluster analysis was used to establish which sounds fell into each of the three experimental conceptual categories. Of course, hierarchical cluster analysis was not necessary to determine membership to the prototype category.

Just as in Chipley (2004), the purpose of the described set of experiments was to explore the boundaries of the sound level cue to distance. How qualitatively-different sounds and qualitatively-similar sounds are judged in distance, relative to each other, when there are changes in sound level is important to a full understanding of auditory distance perception. Specifically, would one 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 experiments, we might predict an affirmative answer to such a question.

To be clear, the following were specific hypotheses regarding the outcomes of the third and fourth experiments:

H₁ – Decreases (increases) in level will be associated with increases (decreases) in judged distance.

H₂ – Sounds which are otherwise completely matched will have the greatest judged distance between them with increases in relative sound level difference.

H₃ – Sounds that fall within the “same” conceptual category should have greater judged distance between them with increases in relative sound level difference than sounds that fall within "similar" conceptual categories.

H₄ – Sounds that fall within a "similar" conceptual category should have greater judged distance between them with increases in relative sound level difference than sounds that fall within “different” conceptual categories.

General Overview

Four experiments were necessary, in order to adequately address the aforementioned hypotheses. The first experiment asked people to verbally identify a representative sub-set (75 sounds) of the 120 normalized sounds created, in part, and edited by Marcell et al. (2000). Further details of the stimuli are discussed later. Experiments 1 and 2 took place in the office space directly adjacent to the audition lab. Sounds were presented with a Dell® OptiPlex GX 400 Personal Computer through Koss® UR29 closed-ear headphones. For Experiments 3 and 4, sounds were presented in the audition lab space previously described in Chipley (2004) and described again below.

The series of experiments investigated how the similarity of familiar environmental sounds affected judgments of perceived relative distance.

Experiment 1

Method

Participants. Experiment 1 involved 5 men and 5 women. They ranged in age from 18 to 24 and participated in partial fulfillment of a research requirement for their introductory psychology course. 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 (1992) was used to assess overall hearing level, and was used as a criterion for participation in the experiment. Participants who failed to score a satisfactory score on the HSI were replaced. 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 (1992), the HSI was correlated with pure-tone hearing thresholds. The inventory had an internal consistency coefficient of .89 and test-retest stability coefficient of .88. Classification of severity of hearing loss with HSI matched that of standard audiometric testing 92.1% of the time.

Stimuli. Table 1 shows the set of 120 normalized sounds (used by Marcell et al., 2000) from which the current stimuli were taken. Marcell took his sounds from a number of sources, including past studies (80 sounds), sound effects CDs, and recordings of everyday sounds (40 sounds). Although Marcell used some of the sounds unedited, most were changed in some way, including reducing duration, increasing duration, reducing or

increasing sound level, removing extraneous silences and noises, and applying a fade-in and fade-out algorithm to the beginning and end of each sound.

Marcell saved each sound as a 16 bit .WAV file at a sampling rate of 22,050 Hz. He did not standardize sound duration. He made this choice, because – in keeping with an ecological stance on environmental sounds—sounds in the real world will not always have the same duration. To be specific, the sounds had a mean duration of 2406 ms (SD 1306 ms) and a range of 137 ms to 6083 ms. (There was no correlation between the duration of the sounds used and either the accuracy of the identification of the sounds or the confidence in such identification. Thus, the choice seems to have been sensible.)

Marcell's 120 sounds were available online for download at

<http://www.cofc.edu/~marcellm/confrontation%20sound%20naming/zipped.htm>. They were downloaded and used unaltered in the current study.

Table 1: Normalized environmental sounds from Marcell et al. (2000).

1. Telephone	41. Gargling	81. Guitar
2. Toilet flushing	42. Horse galloping	82. Jackhammer
3. Birds chirping	43. Lawn mower	83. Crow
4. Brushing teeth	44. Sawing	84. Woodpecker
5. Sneeze	45. Snoring	85. Accordion
6. Church bells	46. Thunder	86. Hammering
7. Clearing throat	47. Trumpet	87. Typewriter
8. Frog	48. Wind	88. Water draining
9. Knocking	49. Crickets	89. Banjo
10. Airplane	50. Organ	90. Gunshots
11. Baby crying	51. Cymbals	91. Ping-Pong
12. Car horn	52. Mosquito	92. Harp
13. Laughing	53. Ocean	93. Crumpling paper
14. Whistling (lips)	54. Piano	94. Donkey
15. Cow	55. Truck	95. Car opening
16. Doorbell	56. Chickens	96. Pig
17. Police siren	57. Scream	97. Chewing
18. Pouring water	58. Sheep	98. Sandpaper
19. Camera	59. Rain	99. Gong
20. Cuckoo clock	60. Turning pages	100. Monkey
21. Whistle (instrument)	61. Boat horn	101. Firecracker
22. Yawning	62. Helicopter	102. Water dripping
23. Cat	63. Wind chimes	103. Wolf
24. Dog barking	64. Water bubbling	104. Bongos
25. Rooster	65. Tea kettle	105. Drill
26. Burp	66. Owl	106. Seal
27. Door closing	67. Violin	107. Swords
28. Harmonica	68. Basketball	108. Cutting paper
29. Glass breaking	69. Blowing nose	109. Cork popping
30. Saxophone	70. River	110. Blinds
31. Bagpipes	71. Zipper	111. Explosion
32. Motorcycle	72. Frying food	112. Rattlesnake
33. Train	73. Flute	113. Velcro
34. Cash register	74. Duck	114. Pullchain lightswitch
35. Bowling	75. Lion	115. Stapler
36. Child coughing	76. Car crash	116. Can crush
37. Dropping ice in a glass	77. Clapping	117. Whip
38. Coin dropping	78. Shuffling cards	118. Pinball
39. Drums	79. Machine gun	119. Tearing paper
40. Elephant	80. Bicycle bell	120. Sonar

Within the aforementioned limited set, those sounds which – within themselves – indicated changes in distance, like a motorcycle driving by or an airplane landing, were not used. This was important, because the primary task in Experiment 4 was to judge the relative distances of pairs of sounds. Having to judge the change in distance represented within one or both of those sounds individually would have been too difficult a task and would have confounded the issues in which we were interested. Considering the fact that there are no data on duration and distance judgments, stimuli for the described set of experiments was limited such that the longest sound was no more than five times as long as the shortest. After removals due to duration and internal indications of distance, 75 sounds remained. Those 75 sounds had a range of 581 ms to 2896 ms, with a mean duration of 1733 ms. See Table 2 for the list of sounds that were used in Experiment 1.

Procedure. Participants were asked to identify each sound with which they were presented via earphones by typing in their response next to the link for that sound in an Excel spreadsheet. Hyperlinks were created in an Excel spreadsheet so that listeners could simply click and listen. The links were labeled in a non-descript way (e.g. “Sound 1”, “Sound 2”, etc.), in order not to give any hints about a sound’s identity. The sound links were randomly ordered for each participant. The participants were simply instructed to “use one or two words to describe what you hear.”

Results

The first set of data collected in the experiment involved scores from Coren and Hakstian’s Hearing Screening Inventory (HSI). The HSI scores suggested that none of the listeners were experiencing hearing loss.

Table 2: The 75 normalized environmental sounds taken from Marcell et al. (2000).

1. Stapler	26. Sandpaper	51. Clapping
2. Sneeze	27. Frog	52. Rooster
3. Cat	28. Seal	53. Coin dropping
4. Water dripping	29. Bongos	54. Sonar
5. Tearing paper	30. Piano	55. Blinds
6. Whistle (instrument)	31. Whistling (lips)	56. Explosion
7. Dog barking	32. Blowing nose	57. Lion
8. Car horn	33. Telephone	58. Boat horn
9. Scream	34. Owl	59. Organ
10. Camera	35. Knocking	60. Crow
11. Sheep	36. Trumpet	61. Hammering
12. Cow	37. Mosquito	62. Donkey
13. Drums	38. Doorbell	63. Typewriter
14. Chickens	39. Baby crying	64. Police siren
15. Cymbals	40. Cutting paper	65. Rattlesnake
16. Zipper	41. Birds chirping	66. Gargling
17. Glass breaking	42. Dropping ice in glass	67. Jackhammer
18. Gunshots	43. Swords	68. Monkey
19. Clearing throat	44. Machine gun	69. Yawning
20. Crickets	45. Harp	70. Basketball
21. Can crush	46. Laughing	71. Bagpipes
22. Whip	47. Sawing	72. Drill
23. Velcro	48. Wind chimes	73. Car crash
24. Woodpecker	49. Crumpling paper	74. Child coughing
25. Bicycle bell	50. Firecrackers	75. Duck

Marcell et al. (2000) gave very specific directions regarding how verbal reports are to be judged in terms of accuracy of identification, and those guidelines (found in Appendix B) were followed. Only the sounds that were the most accurately identified (as measured by percentage of listeners who correctly identified the sound) in the first experiment were used in the experiments that followed. Only sounds that were correctly identified at least 80% of the time were used. This criterion left 55 of the original 75 sounds for use in Experiment 2.

Table 3: The 55 sounds yielded from the original 75 sounds. These were the sounds that were correctly identified at least 80% of the time by listeners in Experiment 1.

1. Blowing nose	29. Camera
2. Chickens	30. Car crash
3. Crow	31. Cat
4. Crumpling paper	32. Child coughing
5. Donkey	33. Clearing throat
6. Duck	34. Coin dropping
7. Firecrackers	35. Cow
8. Glass breaking	36. Crickets
9. Harp	37. Doorbell
10. Monkey	38. Drums
11. Mosquito	39. Explosion
12. Sonar	40. Frog
13. Woodpecker	41. Gargling
14. Baby crying	42. Gunshots
15. Basketball	43. Knocking
16. Clapping	44. Laughing
17. Cymbals	45. Piano
18. Dog barking	46. Rooster
19. Hammering	47. Sawing
20. Lion	48. Scream
21. Organ	49. Sheep
22. Owl	50. Sneeze
23. Police siren	51. Telephone
24. Seal	52. Trumpet
25. Swords	53. Whistle (instrument)
26. Wind chimes	54. Whistling (lips)
27. Zipper	55. Yawning
28. Birds chirping	

Discussion

It was crucial to verify that experimental sounds had clear identifications, because similarity depended on it. In order to judge the semantic/conceptual similarity of two sounds, one must be able to, at the very least, identify each of the sounds. Experiment 1 yielded 55 reliably identified sounds. These sounds were then categorized in Experiment 2.

Experiment 2

Method

Participants. Experiment 2 consisted of 20 people (10 men and 10 women). Like the participants for Experiment 1, and the participants for Experiments 3 and 4, they were undergraduate students who participated in partial fulfillment of a research requirement for their introductory psychology course. Normal hearing was required, as assessed by the HSI. The participants in this experiment ranged in age between 18 and 29; all met the established criterion.

Environment and apparatus. The environment was the same as for Experiment 1. The apparatus included a deck of 2" X 3" note cards with printed labels of the 55 sounds obtained from exp. 1. Marcell's (2000) labels for the sounds were used.

Stimuli. Experiment 2 involved the presentation of a deck of cards with the written labels.

Procedure. This experiment involved the classification of non-heard environmental sounds into different conceptual categories. How participants classified the sounds was up to them. A sorting task to classify sounds into different conceptual groups, similar to the procedure used in Bonebright (2001) was used. Participants were given a deck of cards with sound-event labels on each card (e.g., "crumpling paper", "car crash", "clearing throat"). Participants were instructed to sort the cards into different piles. Each of the piles was to be composed of what the participants believed to be conceptually similar sound events. Examples were given verbally to be sure that they understood the task. Specifically, "Mars", "Jupiter", "Bear", "Dog", "Godzilla", and

“Bigfoot” were given as examples of things that might be considered conceptually same, similar, or different. “Mars” and “Jupiter”, both planets, would be considered very similar, if not the same, conceptually. “Bear” and “Jupiter” would likely be considered as conceptually different, with the first being an animal and the second a planet. Finally, “Bear” and “Bigfoot”, might be considered conceptually similar, with both being animals, but only one imaginary. Participants were instructed to focus on, and make their judgments based on the *conceptual* similarity of indicated events, and not upon how they imagined such events would sound. No constraints were placed on how many groups could be created. The only rule was that there had to be at least two sounds in each group. After the card sorting was completed, the experimenter recorded which sounds were put into each pile. The participants were then asked to combine his/ her piles, in order to create higher-level groups. In other words, some of the piles that a participant created may have been similar to one another and may thus have belonged together in “groups of groups.” The experimenter also recorded which sound groups were combined. This additional step allowed the experimenter to identify the sounds that fell into “similar” conceptual categories.

Because of the need for participants to understand both the written sound labels on each of the cards and the specific instructions with regard to how the cards should be grouped, only native speakers of English were used. Language and cultural differences could also have had an effect on how sounds were grouped into conceptual categories.

Design. The experimental design was mixed. Every participant was exposed to all stimuli (within-subjects). The only between-groups factor was gender.

Results

Hierarchical cluster analysis was used to determine which sounds were conceptually grouped together. Specifically, a hierarchical tree cluster (a graphical structure that represents relationships between stimuli) was created. This structure, sometimes called a dendrogram, was created using a software package specifically designed for card-sorting tasks. The program was created by IBM ® and is called USort™. USort is accompanied by a sister program called EZCalc™. USort was used to enter the original list of sounds presented to each participant. Then, the first groupings of sounds, as indicated by the card-sorting of the participants was entered, followed by the second higher-order groupings of sounds, also indicated by the card-sorting of the participants. This was done for each participant. Once all of the card-sorting data was entered into USort, EZCalc was used to create a dendrogram to summarize the results.

For the purposes of describing a dendrogram and the information that may be gleaned from one, reference to Figure 1 may prove helpful. The sample dendrogram has a number of branches that come together at nodes or connections, and represents data from an imaginary group of participants. Branches that share the same node are part of the same group. Furthermore, Euclidean distance (indicated on the X-axis of the dendrogram) can be used to quantify the separation between nodes. For the current research, this separation on the X-axis aided in the identification of separate clusters or categories, as in "same," "similar," and "different" conceptual categories of sounds. The different colored lines or “branches” indicate low-level groupings. The alternating

orange and blue color blocks indicate higher-level groupings. The vertical green and pink lines indicate criteria flags to indicate cut-offs for group membership. The default settings are at .3 and .7 respectively. The user may adjust the position of these flags if desired, and the values within the green and pink boxes change as the user changes the corresponding lines' positions.

In order to render a single dendrogram from all of the sorting data of all of the participants, an amalgamation algorithm was used. USort has three built-in algorithms that can be used to amalgamate data: a single-linkage algorithm (which emphasizes similarities across participants), a complete-linkage algorithm (which emphasizes differences across participants), and an average-linkage algorithm (which gives a good balance of the previous two). Since this study was concerned with similarities *and* differences, the average-linkage algorithm was utilized.

Choosing three relatively equally-spaced clusters that comprised the three conceptual categories of sounds was a somewhat subjective process. However, the dendrogram made the task easier by providing a visual basis for making decisions. Sounds that shared a primary node were labeled as “same” conceptual category sounds. Sounds that shared a secondary node were labeled as “similar” conceptual category sounds. Choosing "similar" sounds required some decisions, but sounds that shared a node that fell somewhere between .3 and .7 were sought out, because this would indicate sounds that fell into a grouping somewhere between lower level (“same”) and higher level (“different”), although this didn’t always happen. Finally, sounds that only shared the highest level node (i.e., the node that indicates that the only thing that sounds

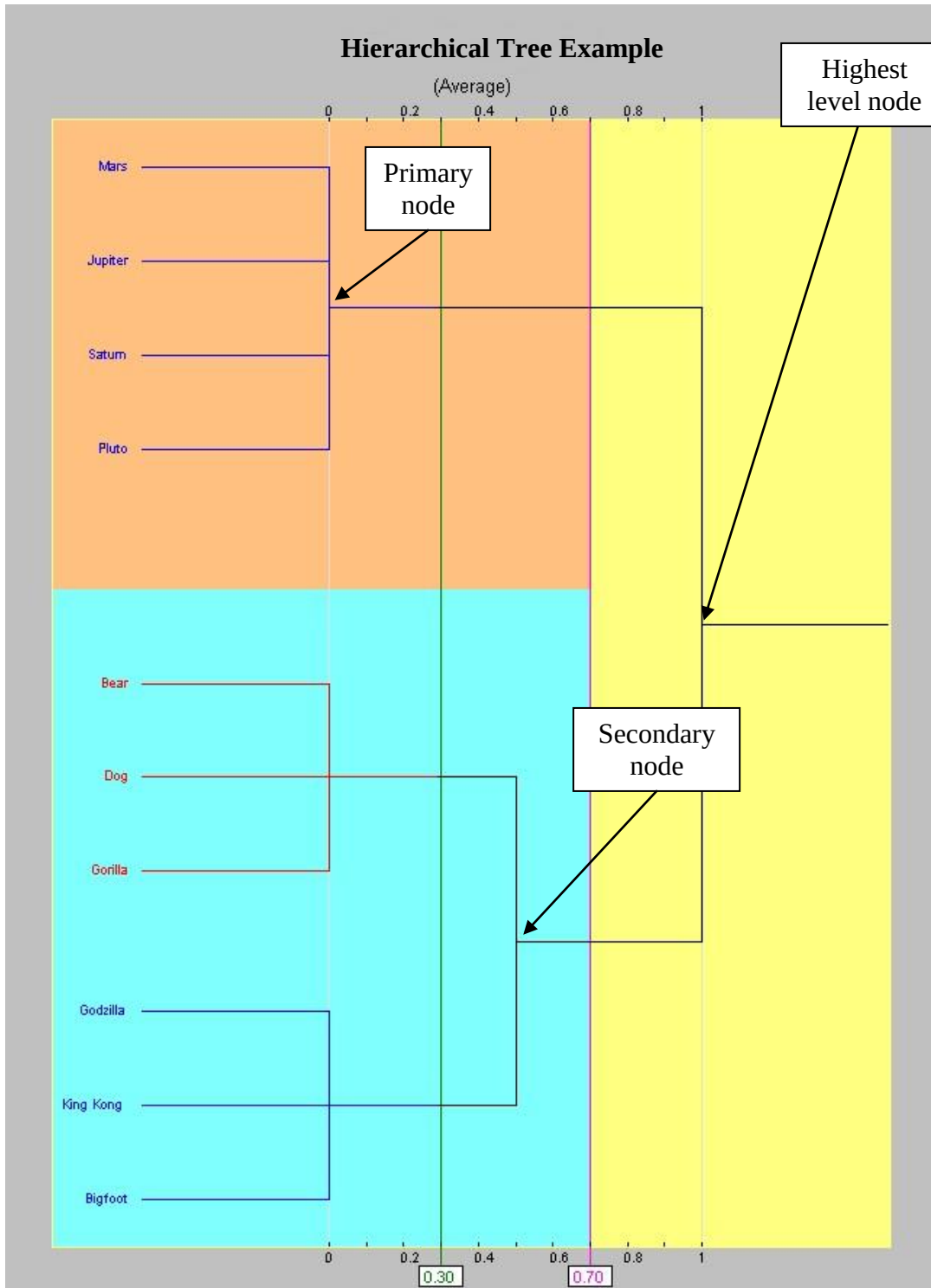


Figure 1: Example of a dendrogram.

converging there have in common is that they are all sounds were be labeled as “different” conceptual category sounds. These were sounds that only shared the nodes at a Euclidean distance between .8 and 1 on the dendrogram. For example, in the sample dendrogram in Figure 1, dog and bear would fall into the “same” conceptual category. Dog and Bigfoot would fall into “similar” conceptual categories. Dog and Jupiter would fall into “different” conceptual categories.

Following the aforementioned process, a total of 18 of the 55 sounds were chosen that best met the requirements of falling into the “same,” “similar,” or “different” categories (see Figure 2). The goal was to identify 6 sounds that fell into each category. Instead, 6 fell into the same category, 7 fell into the similar category (albeit with different levels of similarity), and 5 fell into the different category (see Table 4).

As I mentioned before, choosing sounds that fell into different conceptual categories was subjective. For the current set of sounds, and the resultant groupings, sounds that might not be identified as wholly separate sounds (even if they fell into the same conceptual category, like chickens and roosters) had to be avoided. The sounds “clearing throat” and “gargling” is another good example, as well as “whistling (the instrument)”, and “whistling (of a human).” There was no experiment to determine the absolute discriminability of the sounds, but pairs of sounds that might have been readily confused were intentionally excluded. Also, startling sounds like breaking glass, explosions, and gunshots were not used. Yet another factor that had to be considered was using all animal sounds, human sounds, or environmental sounds. While a fair balance was desired, given the self-imposed constraints, a perfect balance proved to be impossible.

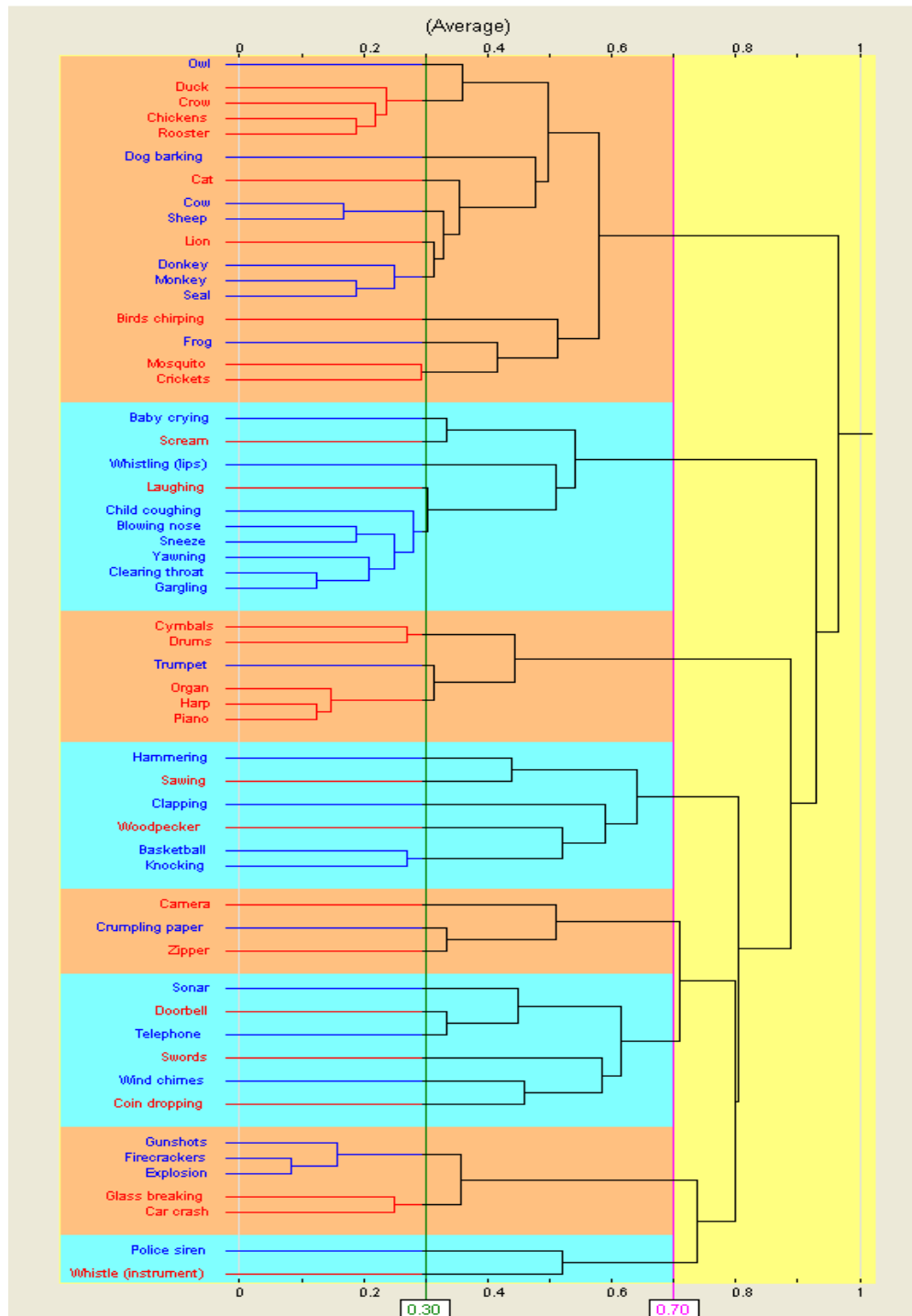


Figure 2: Dendrogram indicating the resultant clustering of sounds.

Table 4: Sounds were identified from the original 55 that best fell into same, similar, and different categories. Six for same, 7 for similar, and 5 for different.

Group	Sound Label	Conceptual Category
1	Cow Donkey Harp	Same Similar Different
2	Sheep Frog Whistling	Same Similar Different
3	Crow Dog Bark Doorbell	Same Similar Different
4	Rooster Owl Windchime	Same Similar Different
5	Child Coughing Baby Cry Crumpling Paper	Same Similar Different
6	Clearing Throat Sawing Knocking	Same Similar Similar

Discussion

Actual sounds were not used, in order to minimize the influence of the acoustic aspects of the sounds. However, it is likely that participants' judgments were biased, in part, by their memory of particular sounds or how they imagined that the event would actually sound. It is also important to address the possibility, and perhaps the likelihood, that how sounds were grouped depended somewhat on the choices provided (i.e., the

particular set of 55 sounds provided for sorting). In other words, it is likely that there were some context effects.

For example, if Saturn, Jupiter, and Pluto were part of a larger list of things to be grouped, they would probably be grouped together on account of the fact that they are all planets. However, if Pluto were the only planet in the list and other things like Disney, dog, and bone, were part of the larger list of items to be grouped, Pluto would likely be grouped with Disney, dog, and bone. Although there are methods to address such a problem (Tversky, 1977), the effect was more or less unavoidable in this set of experiments.

Although a multidimensional scaling approach could have been used to assess how underlying physical aspects of the sounds affected why people grouped things together, this experiment was not concerned with *why* things were grouped together as much as with the fact that things either *were* being grouped together or *were not* being grouped together. “Why” is a good question, but beyond the scope of this investigation. To answer the question of whether or not conceptual similarity of sounds affects perceived relative distance between them, we only needed to know what sounds were grouped together or separately, and to what degree.

A card-sorting approach seemed like the most efficient, given the number of sounds that had to be used. Studies using visual and tactile stimuli often use sorting tasks (Garbin, 1988). Most importantly, the validity of using sorting tasks for sounds has been tested and shown to be effective with two different sets of auditory stimuli (Bonebright, 1996). Also, a paired-comparison task would have been extremely cumbersome. The

ultimate success of the entire set of experiments hinged on having groups of sounds that were reliably identified as coming from conceptually same, similar, or different categories. Although the resulting selections were not quite as diverse (and clearly separated) as originally expected, Experiment 2 provided the necessary groups to continue the research.

Experiment 3

Method

Participants. Experiment 3 consisted of 36 people (18 men and 18 women). The age range was between 18 and 33.

Rationale

The third experiment involved the presentation of a smaller subset of sounds from the same, similar, and different categories established by Experiment 2. This experiment served to "equate" the sounds to be presented in the fourth and final experiment. "Equate" is used loosely here. The term does not imply calibration, but simply identifying those sounds that seem to have similar sound level vs. perceived distance relationships. Thus, measures such as slope and intercept for each function were compared between sounds. It is important to understand why this step was necessary. Consider a high-pitched bell producing a 70 dB level at the listener's head. There is no *a priori* reason that the bell should seem to be at the same distance as a 70 dB level produced by a dog barking. Similarly, the relationships between sound level and perceived distance for each of the previously mentioned sounds are not likely to be the same. There could be a number of reasons for such an outcome, including, but not

limited to, differences in spectral characteristics, importance to the listener and the “normal” acoustic output of the sounds presented (i.e., familiarity). Since the goal of the main experiment (Experiment 4) was to assess the effect of sound similarity on the usefulness of cues to perceived relative distance, it was important to use only those sounds that, for a given sound level, were heard at roughly the same distance and for which a given change in sound level resulted in a similar change in perceived distance.

Environment and apparatus. Experiment 3 took place in a laboratory consisting of an acoustically treated test room and an adjacent control room where sounds were generated and manipulated using the aforementioned computer. The test room contained a keyboard and monitor connected to the laboratory computer, so that stimuli could be activated remotely. 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, keyboard and monitor, speaker and assorted other equipment. All presentations were completed in near darkness. Listeners were only able to see the table at which they were seated and some dark-colored curtains in front of them.

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. Overall, the room had reverberation characteristics that made it acoustically “dead.”

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 cup 94 cm above the floor. The height of the chin cup placed the average person's ears at the same height as the speaker used to present sounds (102 cm above the floor). The experimenter instructed the listener and collected data from a position behind another curtain, 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 heavy cardboard tube. The support structure and speaker were physically located 2.5 m in front of the listener.

Stimuli. Each of the 18 sounds (as determined by Experiment 2) from each category (same, similar, and different) was randomly presented at 10 different sound levels (from 30 to 84 dBA in 6 dB increments). This resulted in a total of 180 sounds. Not wanting to present everyone with each sound, the sounds were broken down into 6 groups of thirty sounds each. The groups of sounds were the same as those shown in Table 4.

Sound level was measured in decibels (dB) A-weighted, or dBA. The A-weighted decibel is the most common unit used for measuring environmental sound levels. It adjusts, or weights, the frequency components of sound to conform to the mid-level response of the healthy human ear. Specifically, the lower frequency components of a sound are weighted less. Sound level varied randomly within each of the presentation groups described in Table 4. An Excel spreadsheet was used to generate a random order

of presentation of stimuli for each participant. The inter-stimulus interval (ISI) for participants was approximately 10 – 15 seconds. Such a short ISI may have led to order effects within participants, but random presentation between participants should have “washed out” any such order effects, thus minimizing the effects of relative cues in this experiment. Each participant wore earmuffs for one minute prior to the first sound presentation to minimize any possible threshold shifts as a result of noise exposure prior to experimentation.

Cool Edit Pro ©, a software application used for generating, manipulating, and analyzing sounds, was used to equalize RMS values across the sounds, so that all sounds could be presented at as close to the same levels as possible. RMS, the acronym for *root mean square*, refers to the square root of the mean of the moment-to-moment amplitudes of a sound squared. RMS was used, because it is a useful representative value based upon a standard acoustical procedure, even when a sound contains internal fluctuations. A sound level meter (Rion NA-61) set on slow response, and positioned at the normal location of the participant’s head, was used as a double-check of sound-level output of the software application. The sound level meter was only a crude check, because of the varied duration and irregular nature of some the sounds being used.

Response measures. Each participant was asked to give a verbal judgment of the distance of each sound event in feet and inches. Other units were also acceptable, as all judgments were converted to meters, an International System of Units base unit.

Design. The experimental design was mixed. Not everyone was exposed to all stimuli. Participants were assigned by order of participation to one of the six stimulus

groups such that three men and three women were exposed to each of the six stimulus sets as detailed in Table 4. Stimulus group was a between-groups factor (though no differences were expected). Gender was another between-groups factor.

There were simply too many presentations (18 sounds X 10 sound levels) to present them all to each participant, in order to have a within-subjects design. The risk in presenting only a few of the sounds to each participant was mitigated by the fact that each participant listened to a sound from each of the three categorical groups, minimizing the chance that a few judgments would affect one category in an unbalanced fashion, relative to the other categories. More importantly, a full 10-level function was generated for each sound within a set of 6 listeners.

Results

The first set of data collected in the experiment involved scores from the aforementioned Hearing Screening Inventory (HSI). Although everyone who participated in the experiment self-reported normal hearing, the HSI scores for Experiment 3 suggested that two men were experiencing some hearing loss. The judged distances of the two 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 median results.

The primary measures for Experiment 3 were median reported distances for each sound at each sound level. Male and female data were combined in the calculation of the medians. The criteria for choosing sounds to use in the final experiment were slope and intercept. Slope, in this case, refers to the degree to which perceived distance (decreases)

increases with (increases) decreases in sound level. Intercept, in this case, refers to the sound level of a sound when perceived to be at a particular distance. For the purposes of the experiment 10 ft. (.30 m) was used as a reference. Sounds with similar slopes and intercepts were desirable for Experiment 4. Thus, for example, if we knew that two sounds required the same sound level to be perceived at a distance of 10 ft. and both sounds' linear models had similar slopes, then we could be confident that, if the systematic increases and decreases in sound level (by 6 dB) in Experiment 4 resulted in differential changes in reported distance, those changes would be due to categorical differences between sounds and not differences in sound models.

No formal statistical analysis was used to choose sounds based on slope and intercept. The experimenter simply looked at the graphs of perceived distance vs. sound level for each of the sounds. There were a few clear rejections (see Figure 3). For individual graphs of median reported distances for each sound see Appendix D.

Many functions fell within a fairly tight grouping of functions. Sounds that fell within that tight grouping were chosen for Experiment 4. The sounds that were chosen were “clearing throat”, “child coughing”, “knocking”, “sheep”, and “whistling (lips)”. Thus, two sounds fell into the “same” category – “clearing throat” and “child coughing” (“S₁” and “S₂” respectively in Experiment 4), two fell into the “different” category – “sheep” and “knocking” (“D₁” and “D₂” respectively in Experiment 4), and just one fell into the “similar” category – “whistling (lips)” (“Sim₁” in Experiment 4).

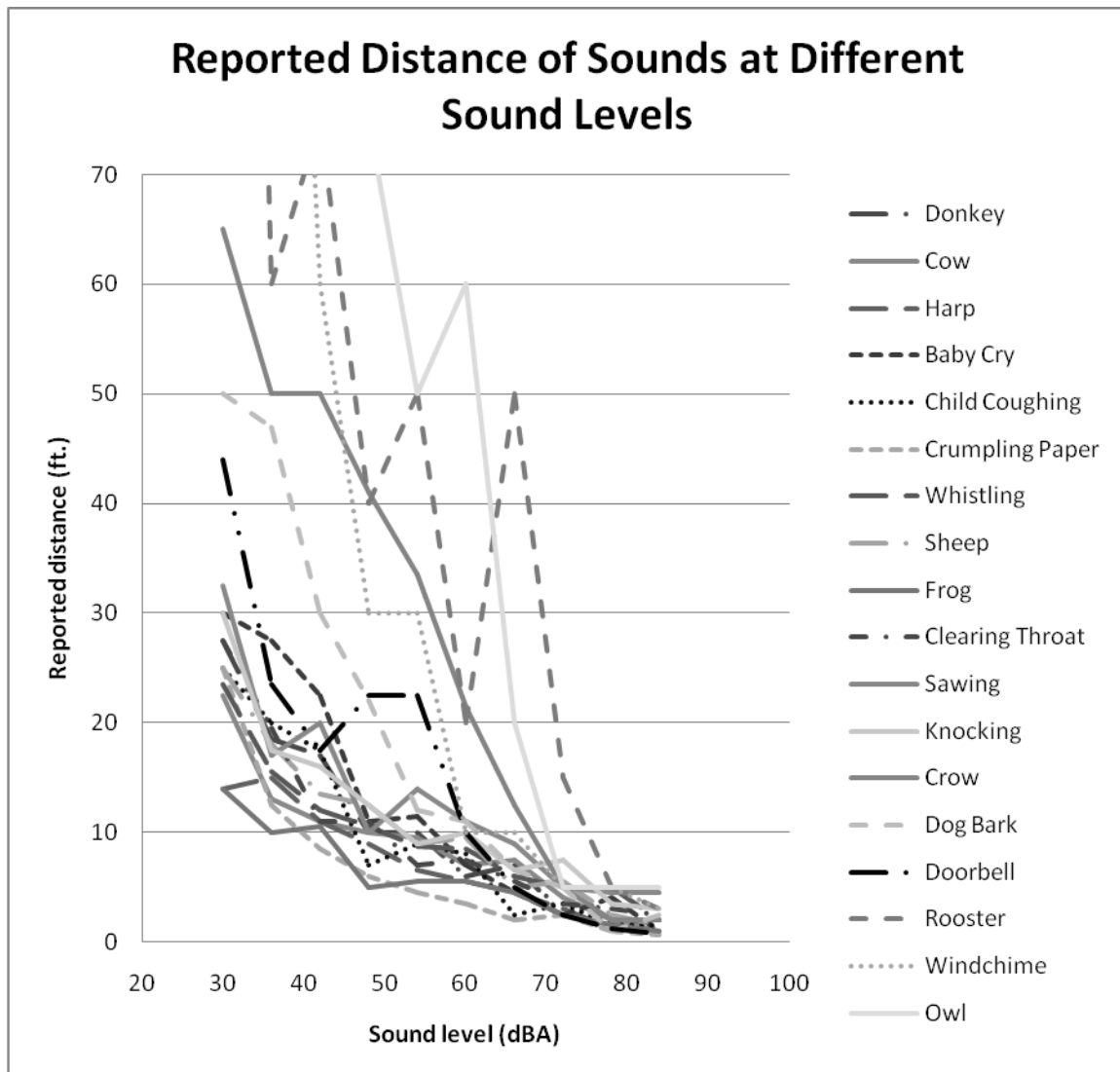


Figure 3: Median reported distances of sounds at different sounds levels. Ten feet was used as a reference intercept for the purposes of comparing the sound functions.

Essentially, the findings of Experiment 3 met the criterion that the sounds chosen for Experiment 4, presented at equal sound levels, should not be perceived at significantly different distances.

Discussion

The sounds that exhibited the most similar slopes and intercepts (as described before) included three “human” sounds, an animal sound, and what is probably best

described as an environmental sound – the knocking (which, of course, might be human-made). The sounds were not only favorable because of their similar sound level, and perceived distance functions, but they were also different acoustically. Acoustic differences were favorable, because any potential conceptual similarity interactions with sound level cues to distance may have been difficult to distinguish from possible acoustic similarity interactions with sound level cues to distance. Two sound events separated in time that sound very much alike and differ in sound level may likely be perceived as simply one sound source changing in distance. The same *may not* be true of two very different sounds separated in time. While the final group of sounds represented a very limited subset of sounds that people might be expected to hear at more or less frequent instances, they were sufficient for an investigation into whether sounds' conceptual similarity would interact with other sound level cues to distance.

Experiment 4

Method

Participants. Experiment 4 consisted of 60 people (30 men and 30 women). The age range was 18 to 36. The HSI scores suggested that a total of six people (4 men and 2 women) were experiencing some hearing loss. The judged distances of the six 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 median results.

Environment and apparatus. The environment and apparatus were the same as for Experiment 3.

Stimuli. For Experiment 4, each stimulus consisted of a sound pair. The first sound (the standard, or reference, sound) in any pair was always the same for a given participant. Two different standards were used in total. Half the participants heard one and half heard the other. The two standards came from within the grouping identified as being conceptually in the "same" category. There should thus have been no differences between groups given that selection of sounds was based on reported (perceived) distance (D') and the changes in perceived distances for changes in sound level ($\Delta d'/\Delta \text{dbA}$) – the functions from Experiment 3. There should also have been no differences if each reference sound really belonged in the "same" category, and if conceptual category affected perceived relative distance.

The second sound of any sound pair was a "comparison" sound. The reference and comparison sounds were separated by an inter-stimulus interval of 500 ms, the same inter-stimulus interval as was used in Chipley (2004). Such a brief period of time between stimuli encourages comparison between sounds in distance - the desired goal. The comparison sound could vary from the reference sound in either sound level or conceptual category or both. Conceptual similarity had four levels and sound level had three. With regard to conceptual similarity, sounds were in prototype, same, similar, or different conceptual categories. The category labeled "prototype" was based on the condition by the same name in the 1993 Coren and Enns study and referred to the condition in which the comparison sound was the exact same as the standard (or reference) sound. With regard to sound level, the second sound in each sound pair was presented at the same sound level as the first (60 dBA), a higher sound level than the first

(+ 6 dB, or 66 dBA), or a lower sound level (- 6 dB or 54 dBA). The plus or minus 6 dB was used to approximate halving or doubling the distances of individual sounds. See Table 5 for the basic design of Experiment 4.

Response measures. Of primary importance, a magnitude-estimation task identical to the one used by Chipley (2004) was used in the experiment. Participants were instructed to consider the standard sound (the first sound in each stimulus pair) to be equal to a distance of “100.” They were asked to always respond to each stimulus-pair presentation with a judgment of the perceived distance of the second sound (the comparison sound) relative to the first. For example, if the second sound seemed to be twice the distance of the first sound, the participant should respond “200.” Each participant made 15 such judgments with an average inter-stimulus interval of 10 – 15 seconds. After the initial 15 judgments of perceived relative distance, each participant was then presented with his/her standard sound alone and asked to judge its egocentric distance verbally in feet and inches. This verbal judgment was used to establish a reference distance by which the indirect judgments of distance could be scaled. For example, if a listener responded “200” for a particular sound pair, and then described the distance of the standard as 5 ft., we could infer that the listener heard the comparison sound at an approximate distance of 10 ft.

The final portion of Experiment 4 involved participants completing a pair-wise comparison task similar in purpose to the card-sorting task done in Experiment 2. Since there were only five sounds, sorting would not have been an appropriate method to assess similarities and differences between sounds. Sound labels were presented in all possible

pairings using an Excel spreadsheet (11 pairings). Participants were asked to judge the *conceptual* similarity of those sound pairings using a Likert scale from 1 (indicating very dissimilar) to 7 (indicating very similar). This was done in an effort to see if the sounds were still grouped in categories commensurate with those in Experiment 2.

Design. The experimental design included 4 factors. The two between-group factors were gender (men vs. women), and reference sound (S_1 vs. S_2). The within-group factors were sound level (same, higher, lower) and category (prototype, same, similar, different 1, different 2).

Gender was an important variable because previous studies have shown sex differences in visuo-spatial abilities (Kimura, 1999) and auditory spatial abilities (Neuhoff & Heckel, 2004). In general, women tend to underestimate distances of auditory looming (approaching) sounds more than men. It has been hypothesized that such a differential has evolved as a safety buffer for women. Whatever the reason, it is important to take sex differences into account in any study of auditory distance perception. Such differences may contribute to a main effect or may interact with other variables.

There also may be confounding factors associated with gender. So, the experimental balance of gender was useful not only to control for “gender differences,” but also to control any differences for which “gender” may serve as a proxy.

Table 5: The design for Experiment 4. The letters and numbers listed under “Group” refer to the sounds that were paired together. The first sound in each pair was the reference sound, and the second sound was the comparison sound. “S” refers to a sound falling in the same conceptual category. “Sim” refers to a sound falling in a similar conceptual category. “D” refers to a sound falling in a different conceptual category. “Type of Comparison” indicates the relationship between the reference and comparison sounds. Within each group, each sound pair was presented 3 times – once with the reference and comparison sounds at the same sound level (same), once with the comparison 6 dB higher than the reference (higher), and once with the comparison 6 dB lower than the reference (lower). There were 15 participants from each gender assigned to each reference (S_1 and S_2) for a total of 60 participants. This table only represents the basic design - not specific presentation orders. Stimulus pairs were randomized with respect to type of comparison (prototype, same, similar, and different) and sound level (same, higher, and lower).

Group 1 (n=30)			Group 2 (n=30)		
Stimuli	Sound Level Relationship	Type of Comparison	Stimuli	Sound Level Relationship	Type of Comparison
$S_1 - S_1$	Same	Prototype	$S_2 - S_2$	Same	Prototype
$S_1 - S_1$	Lower	Prototype	$S_2 - S_2$	Lower	Prototype
$S_1 - S_1$	Higher	Prototype	$S_2 - S_2$	Higher	Prototype
$S_1 - S_2$	Same	Same	$S_2 - S_1$	Same	Same
$S_1 - S_2$	Lower	Same	$S_2 - S_1$	Lower	Same
$S_1 - S_2$	Higher	Same	$S_2 - S_1$	Higher	Same
$S_1 - Sim_1$	Same	Similar	$S_2 - Sim_1$	Same	Similar
$S_1 - Sim_1$	Lower	Similar	$S_2 - Sim_1$	Lower	Similar
$S_1 - Sim_1$	Higher	Similar	$S_2 - Sim_1$	Higher	Similar
$S_1 - D_1$	Same	Different	$S_2 - D_1$	Same	Different
$S_1 - D_1$	Lower	Different	$S_2 - D_1$	Lower	Different
$S_1 - D_1$	Higher	Different	$S_2 - D_1$	Higher	Different
$S_1 - D_2$	Same	Different	$S_2 - D_2$	Same	Different
$S_1 - D_2$	Lower	Different	$S_2 - D_2$	Lower	Different
$S_1 - D_2$	Higher	Different	$S_2 - D_2$	Higher	Different

Results

Method of analysis. A mixed analysis of variance (ANOVA) was used to identify significant differences among conceptual-category treatment groups, sound-level treatment groups, gender effects and interactions between them. Of interest was the effect of changes in sound-level on changes in perceived-distance. Also of interest were any significant interactions between conceptual category and sound level on perceived relative distance between sounds. Positive findings would, most likely, indicate context effects on sound level cues to distance. Given the efforts taken in Experiment 3 to choose sounds with similar sound-level vs. perceived-distance relationships, there should have been no such findings had there been no context effects.

Essentially, the most important question of Experiment 4 was whether there was any interaction between the conceptual relatedness of sounds and the effect of level change on the perceived relative distances of those sounds. As mentioned earlier, it was expected that sounds from the prototype conceptual category would exhibit the strongest depth effect (i.e., the perceived distances among such sounds would be greatest) with changes in sound level. Sounds from the same conceptual category should have had the next strongest depth effect with changes in sound level. Sounds from similar conceptual categories should have had a weaker depth effect than either prototype or same conditions with changes in sound level. Finally, using the same logic, sounds from different conceptual categories should have exhibited the weakest depth effect with changes in sound level.

Sound level. The first hypothesis for the series of experiments predicted that increases (decreases) in sound level should result in decreases (increases) in perceived distance. Such an effect was necessary to demonstrate, in order to conclude that the conditions and sound levels used were such that a strong depth cue was available to listeners. Investigation of possible interactions between other variables (such as type of reference sound or conceptual similarity) and sound level would be moot had we not been able to demonstrate a basic depth effect. A casual look at Figure 4 indicates that changing sound level resulted in changes in perceived distance. Indeed, ANOVA indicated there was a significant main effect of sound level on perceived relative distance ($F(2, 118) = 773.62, p < .0001$).

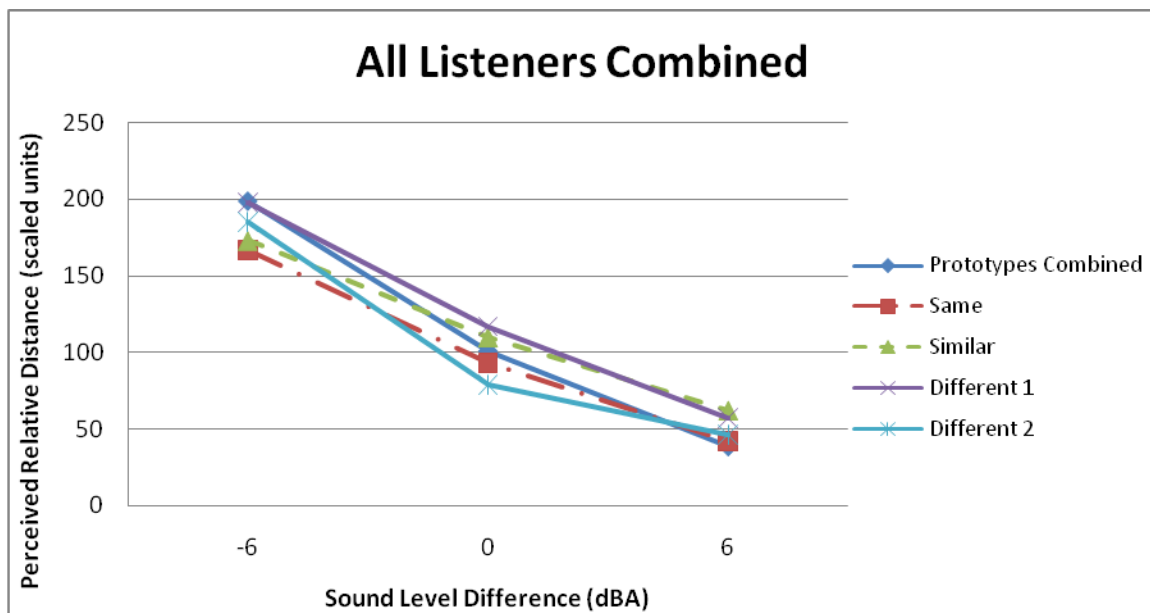


Figure 4: Mean perceived relative distances of the experimental (comparison) sounds at different sound levels. The reference sound was always presented at the same sound level across trials – 60 dBA. Thus, a 0 dBA “Sound Level Difference” indicates a sound level of 60 dBA for the comparison sound. “- 6 dBA” corresponds to a sound level of 54 dBA for the comparison sound. “6 dBA” corresponds to a sound level of 66 dBA for the comparison sound.

Conceptual Category and Sound Level. The second hypothesis of the set of experiments was that depth cue effects would be strongest for sound pairs that were completely matched (i.e., the prototype condition). Remember, the criterion that sounds chosen for Experiment 4, presented at equal sound levels, should not be perceived at significantly different distances was addressed by Experiment 3. The third and fourth hypotheses also involved the relationship between conceptual category and sound level on perceived relative distance. Specifically, sounds that fell within the same conceptual category should have had greater judged distance between them with increases in relative sound level difference than sounds that fell within "similar" conceptual categories. Sounds that fell within a "similar" conceptual category should have had greater judged distance between them with increases in relative sound level difference than sounds that fell within different conceptual categories. Similar sounds, it was reasoned, were more likely to be compared to one another than dissimilar sounds, thus highlighting depth information between such sounds.

There was indeed a significant main effect of conceptual category on perceived relative distance ($F(4, 228) = 16.05, p < .0001$) with conceptual category and sound level also interacting significantly ($F(8, 118) = 4.60, p < .0001$). There was also a significant interaction between reference sound and conceptual category ($F(4, 228) = 6.33, p < .0001$).

Tukey's HSD revealed specific differences in depth cue magnitude between sounds from the various categories. "Different 1" ("sheep") was perceived as significantly farther away from its reference sound than "Different 2" ("knocking").

“Different 1” was also perceived as significantly farther away from both of the “Same” sounds (“clearing throat” and “child coughing”). It’s important to remember that the “Different” category sounds were only in the same category, because they were both judged as conceptually different from both of the reference sounds. They are not similar sounds in comparison with each other (and were not judged to be). Also, “Similar” (“whistling”) was heard as significantly farther away than the “Same” sounds were. “Same”/ “Prototype” (either “clearing throat” or “child coughing” depending on whether they were presented as the reference - “Same”, or with themselves - “Prototype”) and “Different 2” (“knocking”) did not exhibit significantly different $\Delta d'/\Delta dBA$ functions. There was no significant difference in the $\Delta d'/\Delta dBA$ functions between “Same” and “Different 2”.

What all of this means is that there don’t seem to be any systematic differences in the $\Delta d'/\Delta dBA$ functions between sounds falling into different conceptual categories. In other words, the differences in depth cue magnitude seemed not to be systematically influenced by the manipulation of conceptual category. This may be a result of the fact that the sounds in this experiment were specifically chosen to exhibit similar sound level – perceived distance relationships. Thus, one might conclude that it was curious not to find any significant context effects (i.e., systematic effects related to the conceptual categories of the sounds) if there were, in fact, context effects on the $\Delta d'/\Delta dBA$ functions.

Gender and Reference Sound. There was not a significant main effect of gender on $\Delta d'/\Delta dBA$ functions. There was, however, a main effect of reference sound on $\Delta d'/\Delta dBA$ functions ($F(1, 56) = 79.09, p < .0001$). However, surprisingly, there was no

significant interaction between gender and reference sound on $\Delta d'/\Delta dBA$ functions. This was surprising, because the reference sounds themselves were judged as significantly different in distance by verbal report. There was a significant interaction between reference and sound level ($F(4, 118) = 16.05, p < .0001$), as shown in Figures 4 and 5.

Verbal judgments. As mentioned earlier, each participant was also asked to give a verbal judgment of the absolute distance of his/her reference sound as part of Experiment 4. The results are below in Table 6. The results of an analysis of the difference between the judged distances of the reference sounds using a Mann-Whitney U ($U = 51, U' = 47.5, p < .01$) suggest that the perceived absolute (egocentric) distance of Reference 2 for men was significantly lower than that for Reference 1 for men. Mann-Whitney U ($U = 51, U' = 47.5, p < .01$) suggested that the perceived absolute (egocentric) distance of Reference 2 for men was significantly lower than that for Reference 1 for women. Mann-Whitney U ($U = 51, U' = 47.5, p < .01$) suggested that the perceived absolute (egocentric) distance of Reference 2 for men was significantly lower than that for Reference 2 for women.

Table 6: Mean and median judged absolute distance (in meters) of reference sounds 1 and 2 for men and women (with semi-interquartile ranges).

	Men		Women	
	Reference 1	Reference 2	Reference 1	Reference 2
Means	3.59	2.04	3.39	3.01
Medians	3.05	1.83	3.05	3.05
Semi-Interquartile Range	.609	.267	1.18	.914

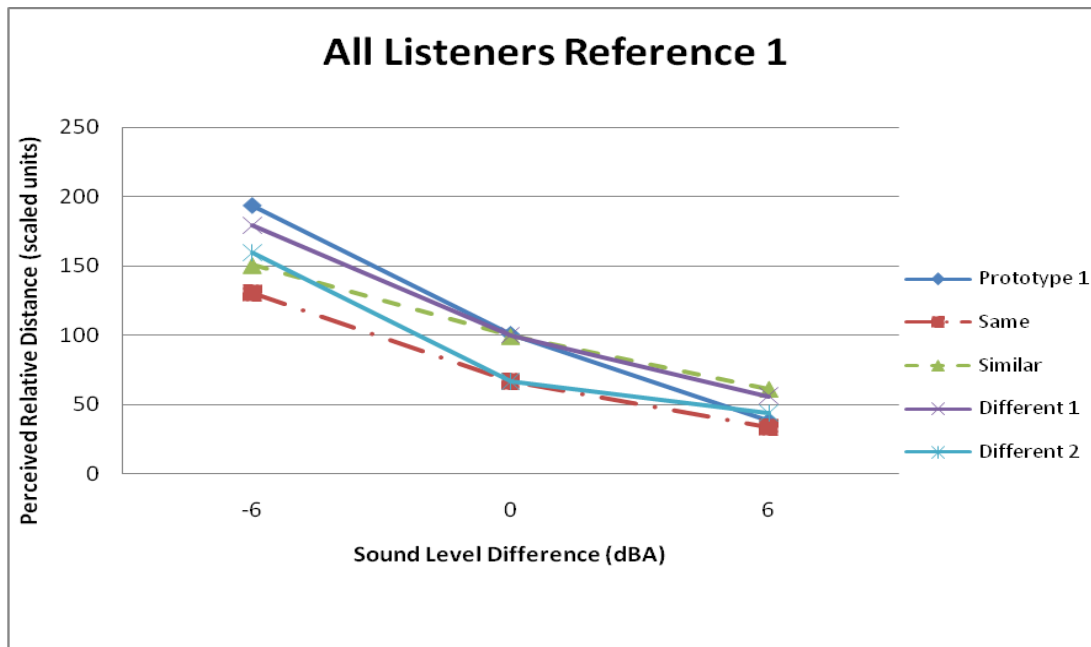


Figure 5: Mean perceived distances of the experimental (comparison) sounds at different sound levels with reference sound 1 ("Clearing throat"). A 0 dBA "Sound Level Difference" indicates a sound level of 60 dBA for the comparison sound.

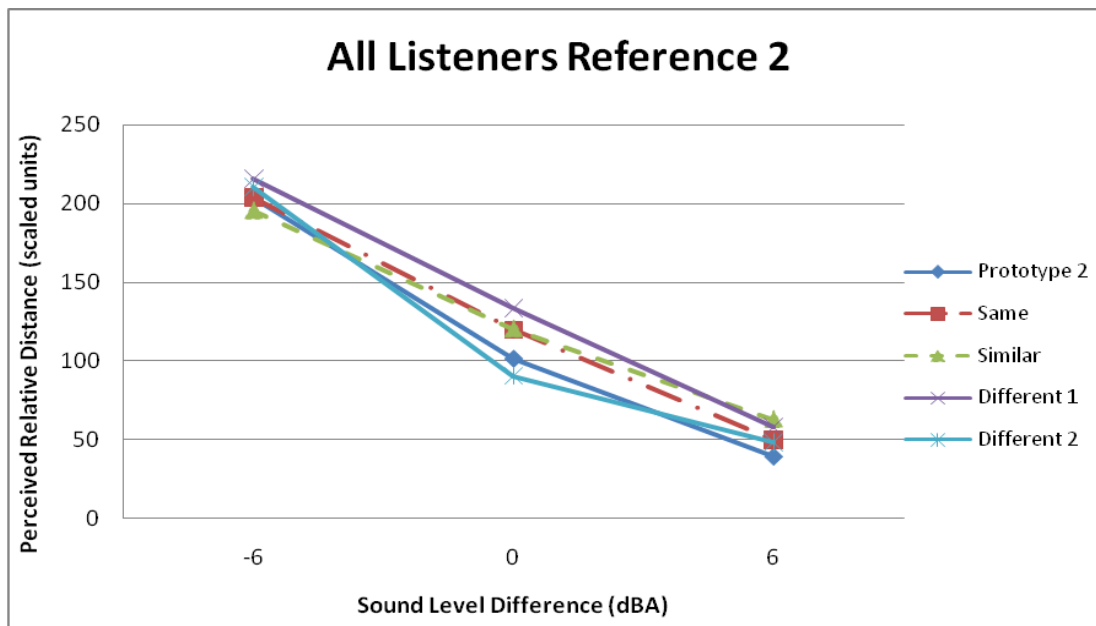


Figure 6: Mean perceived distances of the experimental (comparison) sounds at different sound levels with reference sound 2 ("Child coughing"). A 0 dBA "Sound Level Difference" indicates a sound level of 60 dBA for the comparison sound.

Discussion

There are a few interesting results of the final experiment to consider. These include the apparent lack of any systematic effect of the conceptual similarity of sounds on the effectiveness of relative sound level cues to distance, possible psychoacoustic explanations for the results, the differences in judged distance between the reference sounds, and finally the separate categorization of the experimental sounds of Experiment 4 into conceptual groups, by the participants of that study.

In terms of the effects of conceptual similarity on the $\Delta d'/\Delta dBA$ functions, it may be helpful to draw a parallel between Experiment 4 and a similar study (Chipley, 2004), in which the findings were similar. Specifically, in both investigations, there was no support for the hypothesis that depth cues would be strongest for similar sounds. In the case of Chipley (2004), that meant similarity in terms of envelope and frequency content, whereas in Experiment 4 that meant similarity in conceptual category. Significant interactions between sound level and similarity factors like envelope, frequency content for Chipley (2004), and conceptual category (for Experiment 4) $\Delta d'/\Delta dBA$ functions were found, but the analysis in both experiments provided no conclusive evidence for a systematic effect of the similarity of sounds on the effectiveness of the relative sound level cues. Sometimes dissimilarity in either frequency or envelope or both resulted in increased differences in the $\Delta d'/\Delta dBA$ functions between sounds and sometimes it resulted in decreases in the $\Delta d'/\Delta dBA$ functions between sounds. The same was true for conceptual similarity. Sometimes (in fact, usually) changes in conceptual similarity resulted in no changes in the $\Delta d'/\Delta dBA$ functions.

There is a temptation to probe further for possible psychoacoustic explanations for the results. Without indulging in pure speculation with regard to why the results collected were specifically as they were, a few things may be said of the sounds used in Experiment 4. All sounds were relatively broadband. However, only a few - "whistling", "knocking", and "coughing" - had obvious starts and stops resulting in localization-favorable frequency modulation. "Clearing throat" and "sheep" had comparatively little frequency modulation during presentation. Such differences between sounds likely played some part in their localization in distance. See Figures 7-11 for the spectrograms of the experimental sounds.

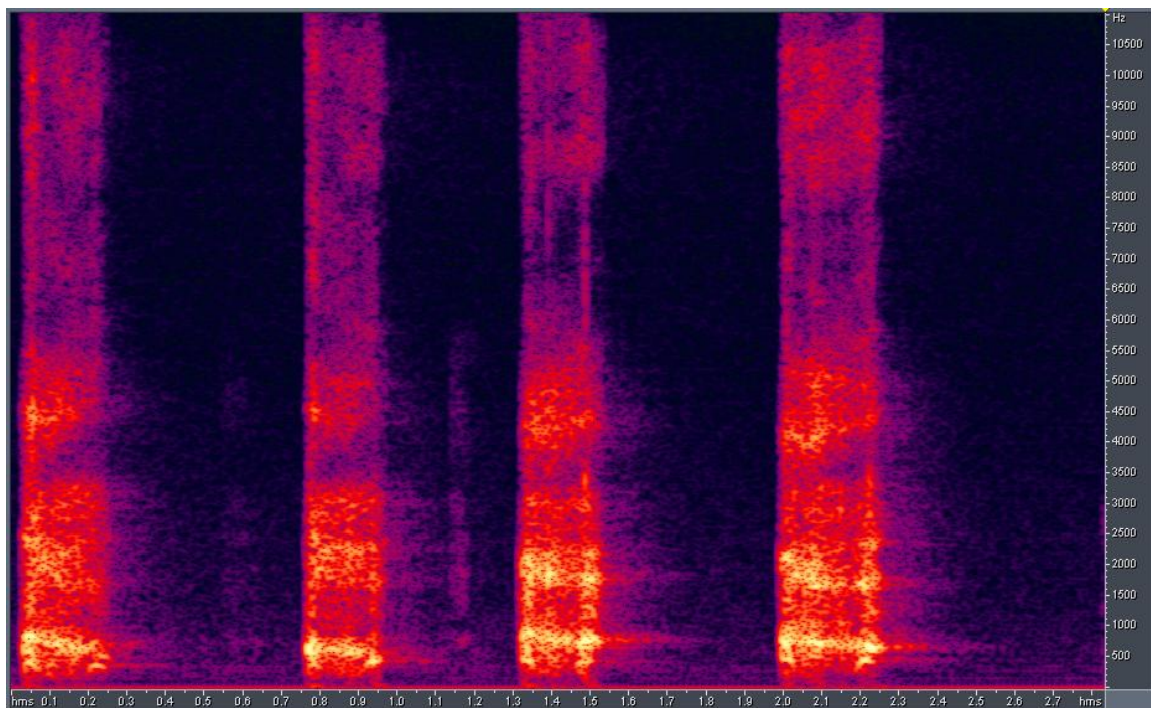


Figure 7: Spectrogram of "Child coughing" - the "Reference 1" sound. Time (sec.) on horizontal axis. Frequency (Hz) on vertical axis.

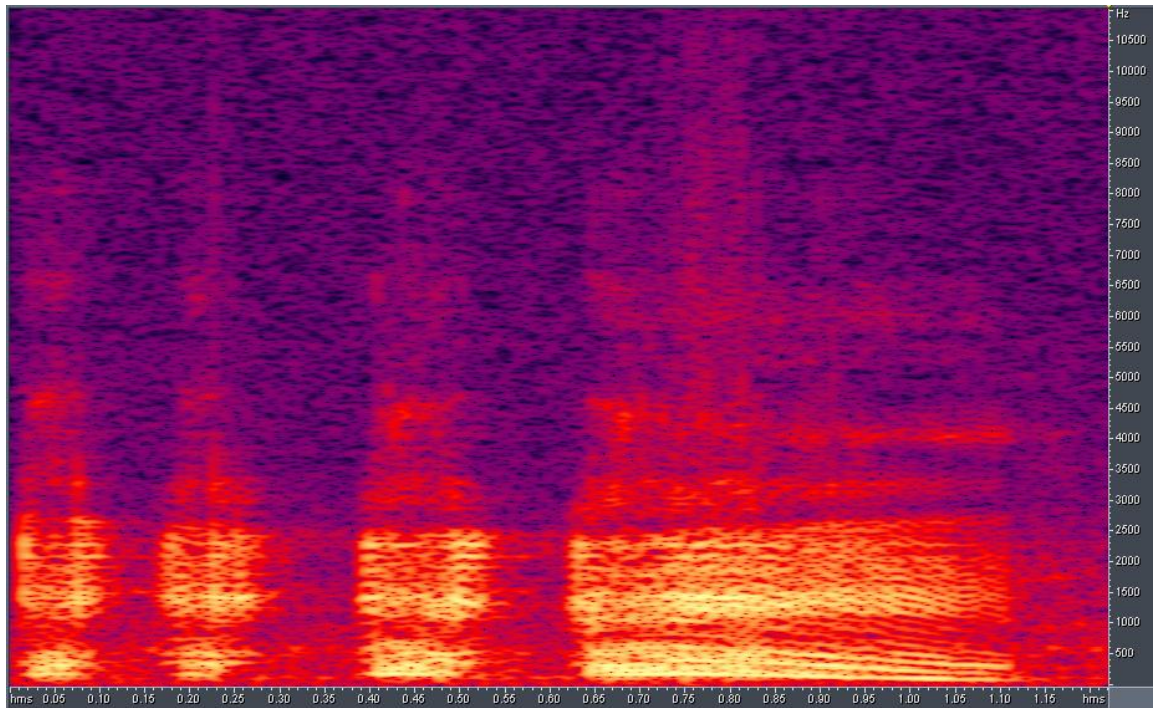


Figure 8: Spectrogram of "Clearing throat" - the "Reference 2" sound. Time (sec.) on horizontal axis. Frequency (Hz) on vertical axis.

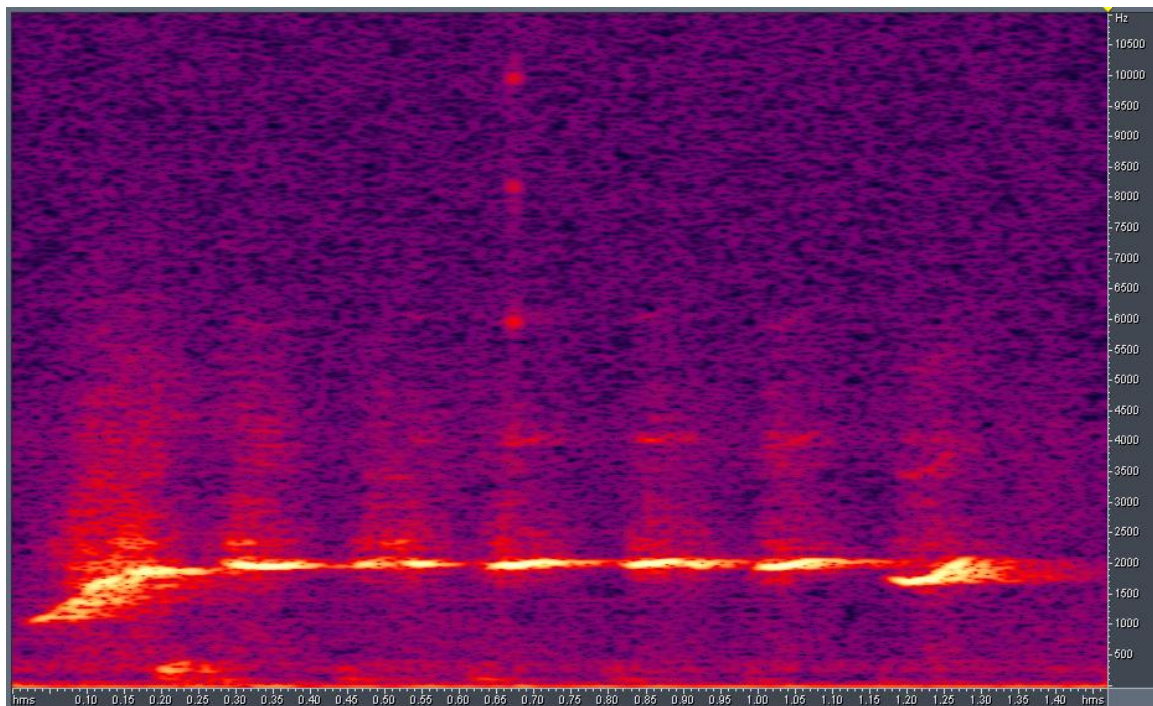


Figure 9: Spectrogram of "Whistling" - the "Similar" sound. Time (sec.) on horizontal axis. Frequency (Hz) on vertical axis.

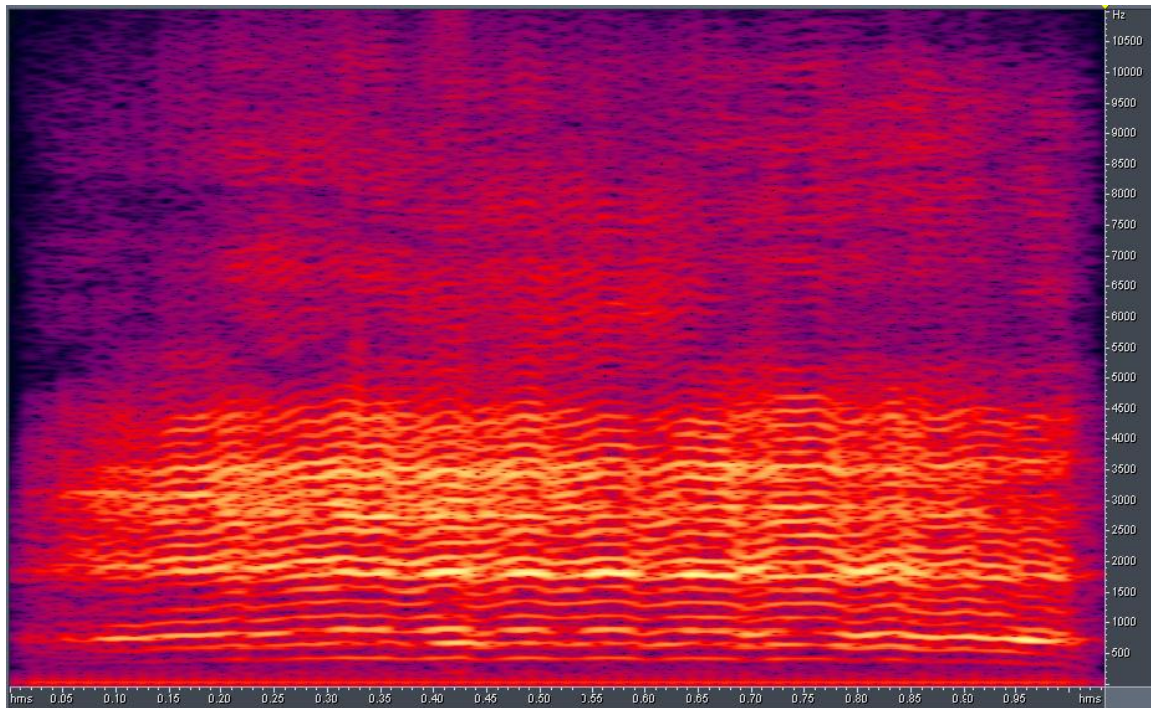


Figure 10: Spectrogram of "Sheep" - the "Different 1" sound. Time (sec.) on horizontal axis. Frequency (Hz) on vertical axis.

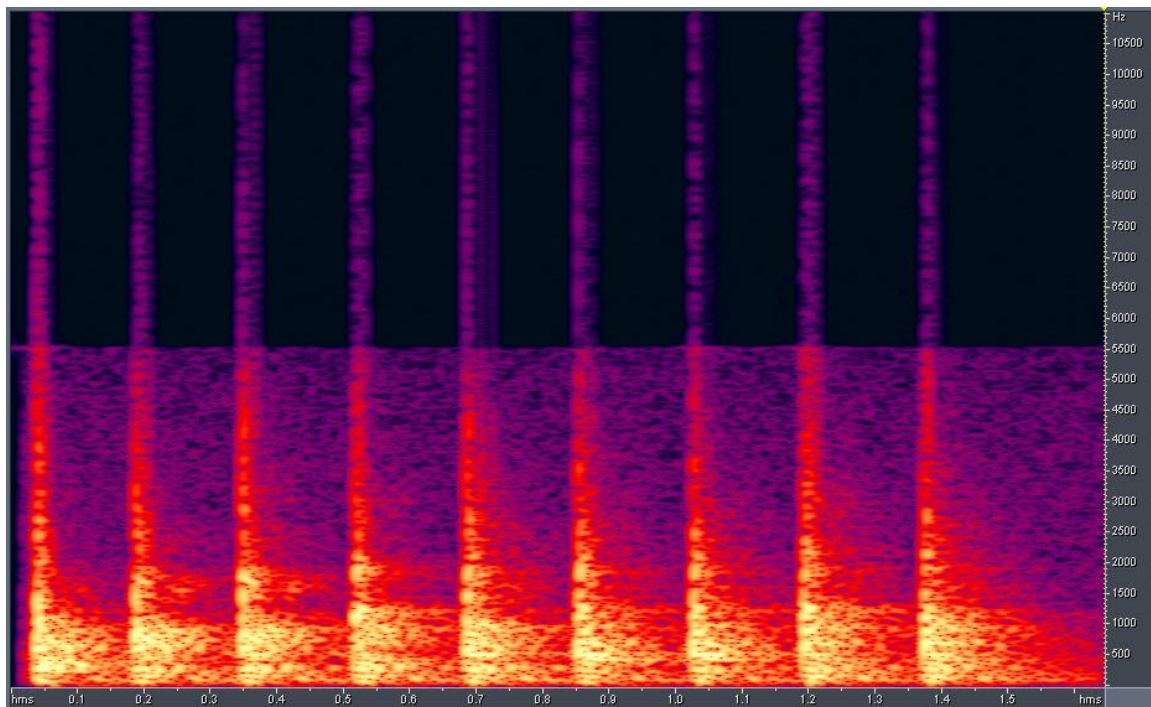


Figure 11: Spectrogram of "Door knocking" - the "Different 2" sound. Time (sec.) on horizontal axis. Frequency (Hz) on vertical axis.

Finally, the differences found between men and women in verbally judged distance of the reference sounds was interesting. First, the finding was interesting, because there were no significant interactions between reference sound and gender on the scaled judgments of reference distance. Also, the finding was curious given the general tendency for women to underestimate distances of sounds more than men. Just the opposite was found here.

Listeners have been found to overestimate distances of nearby sound sources (closer than ~2.75 m) and underestimate distances of more distant ones (farther than ~2.75 m) (Loomis et al., 2002). Thus, given the loudspeaker distance of 2.5 m in this set of experiments, it is unlikely that the physical distance of the sound source played much of a role in the differences found between men and women. With regard to the differences identified between reference sounds 1 and 2, “clearing throat” was about half as long as “child coughing”. Also, as mentioned earlier, “clearing throat” had less obvious stops and starts than “child coughing.” Consequently, it is likely that “child coughing” was perceived differently in distance, because of the strong availability of the onset cue. This is not to imply that there may not have been effects of the familiar sound cue.

Re-conceptualization. The final step in Experiment 4 was to ask listeners to conceptually categorize the 5 experimental sounds, because it could be argued that grouping 5 sounds might be quite different than grouping 60 sounds. The initial thought was, if the 5 experimental sounds for Experiment 4 fell into roughly the same conceptual groups as they did in Experiment 2, we could be fairly confident that there were no

significant context effects and the sounds were still being heard as belonging to same, similar, or different conceptual categories. However, the results turned out to be not that discriminating, and thus, not that helpful for interpretation. The ratings mostly ranged between 3 and 6 with very few people ever using 1 or 7 to describe the similarity of sounds to each other.

A simpler method than described earlier for grouping sounds into conceptual categories was utilized. Specifically, all possible pairings of sounds were presented via an Excel spreadsheet and participants were asked to rate each pairing in terms of similarity from 1 to 7 with 1 being "very dissimilar" and 7 being "very similar". Means were calculated for each sound pairing. The results suggested, though not conclusively (no multi-dimensional procedure was used), that sounds within Experiment 4 were no longer falling into the same conceptual categories as established in Experiment 2. Also, two sounds that should have been rated as very dissimilar (according to the previously established conceptual category of "different") were only judged as relatively similar. All of the other sound pairings seemed to indicate conceptual groupings similar to those in Experiment 2.

The fact that most of the sounds used in Experiment 4 were human-made or animal-made sounds likely contributed to a lack of discrimination. As this measure was taken after the main part of Experiment 4 when people were exposed to the different sounds, their responses were likely influenced by the remembered sounds and the acoustic qualities of those sounds. However, if we may assume that the lack of discrimination (in terms of conceptual category) between sounds existed during the actual

presentation of the sounds, the apparently unsystematic changes in perceived relative distance with manipulation of conceptual category and sound level might not be surprising. This raises an important question. How else are we to ensure that people are hearing sounds as belonging to the distinct conceptual categories that we wish to manipulate? In other words, how do we ensure that the intended manipulation actually occurs, and strongly enough to produce a result?

The people who grouped the sounds into conceptual categories in Experiment 2 were not the same people who eventually judged the respective sounds' distances. It is conceivable that had the same people who grouped the sounds also have been asked to judge their distances; stronger, more systematic results might have been found. However, such a design would have introduced bias into the judgments and sacrificed certain ecological aspects of the current design. The listeners could have been biased by their previous judgments of similarity, and they would have been placed in a situation where they would know exactly what sounds to expect to hear - a decidedly un-ecological situation.

General Discussion

Unlike a similar experiment that preceded this one (Chipley, 2004), real-world sounds (or environmental sounds) were used. The differences in manipulated lab sounds seemed to have had little effect on perceived relative distance. Sounds with manipulated higher-order (conceptual) properties seemed also to demonstrate little systematic effect on sound level cues to distance. While acoustic differences between sounds were not the focus of the present investigation, it might make sense to tie higher-order properties of a

sound to its basic acoustic components. However, since there were no apparent systematic effects on perceived distance with the manipulation of semantic qualities of sounds, there was no real reason to translate such results into an acoustic explanation.

Probably the most important theoretical advance of the present series of experiments involved moving away from a psychoacoustically centered study of human auditory perception. What many in this area of research have come to realize is that ecological investigations of auditory perception are valuable and have the potential to fill gaps in the literature. To that end, the intention of the present study was to use more "real world" stimuli while still maintaining some of the virtues of traditional, tightly-controlled experimentation.

The next logical incarnation of this series of experiments would involve a yet more ecological approach, such as using a more action-based metric for judging perceived distance. At least one recent study (Russell & Schneider, 2006) has shown that using action-based metrics to judge the distance of sound sources (e.g., walking to the sound source) are more accurate than verbal judgments of distance (in feet, inches, meters, etc.) or magnitude estimation. It is argued that laboratory studies of auditory perception generally involve participants performing tasks they normally wouldn't perform in everyday situations (e.g., magnitude estimation). Also, organisms in natural settings are generally always acting or moving. This idea can be traced back to Gibson (1979; 1986) in his coupling of perception and action.

Ashmead, Davis, and Northington (1995) investigated the ability of observers to judge the location of sound sources that varied in distance from 5 to 19 m. Listeners

were asked to walk to the sounds. Perception of distances was nearly veridical. Their study also found that judgments of target distance were less accurate for stationary observers than for observers who were walking during sound presentation. Posture, pivoting, tilting of head, etc. were also allowed. These studies were done outside.

Furthermore, with regard to consistency, Loomis et al. (1998) found listeners to be more consistent in the judgment of target distance when they walked to the sound source versus making a verbal report. Using such an action-based metric could increase precision and perhaps increase the chances of identifying an effect of conceptual similarity on relative cues to distance. It may be argued that the systematic effect of conceptual similarity on relative cues to distance is relatively small if, in fact, it exists at all. Nevertheless, such methodology could increase the chances for detection. However, there are probably other relevant factors, such as whether the final groupings of sounds in Experiment 4 actually held up. In other words, was the *actual* grouping of sounds the *intended* grouping of sounds?

Russell and Schneider (2006) noted that sound-producing objects in natural settings generally vary within a multidimensional space. They pointed to the many investigations into auditory distance perception that show we are typically poor at judging distances of sounds. Of course, such findings do not really jibe with what, anecdotally, seems to be the case in the real world – that we are able to get around and interact with multiple sounds sources without too much trouble, and hear sounds sufficiently well in direction and distance to, in most cases, complete distance-judgment related actions and avoid danger. This suggests that measures of auditory distance

perception used in the past may fail to accurately represent or portray the perceptual abilities of listeners.

However true it may be that many of the metrics used in the past may represent our perceptual abilities poorly, we have to find a happy medium in which sufficient experimental control exists to determine just what factors affect distance judgments. There may very well be simply too many variables to account for in a purely ecological experiment. Consequently, a gradual and measured increase in the balance of ecological methods to traditional experimental methods is recommended here. When traditional methods are unable to capture hypothesized effects of sound/ condition variables in auditory research, slow introduction of ecological methodology is appropriate. However tempting it is to use methods that seem to offer more ecological validity, one must consider the commensurate loss in control that causes internal validity to suffer. Such a deficiency makes it more difficult to rule out alternative explanations for observed effects.

Making verbal judgments of the distances of sounds or using magnitude estimation to judge distances are not the kinds of activities we are used to doing “out in the world,” and may be seen as a weakness of Experiment 4. Another weakness may have been the fact that the sounds were presented in a laboratory which, while listeners could not see the dimensions, could infer that it was probably no bigger in size than most rooms in a typical classroom building. Also, all of the sounds were presented from one speaker that was always at the same distance. However, some of these weaknesses may be seen as strengths in the context of the aforementioned goals of this set of experiments.

As mentioned earlier, one of the goals of the series of experiments was to utilize the virtues of both an experimentally rigorous (and traditional) approach to auditory distance perception research and a more generalizable ecological approach. It made sense to use such a combined approach, in order to allow comparison of results to an experimentally similar previous study (Chipley, 2004).

In closing, there are some "loose ends" that probably deserve some attention. Such things were just outside of the scope of the current experiment. As mentioned before, a more ecological, action- based measure of perceived distance for Experiment 4 might be considered. Aside from, or maybe in addition to, re-running Experiment 4 using a different response measure, a more thorough consideration and dissection of all the acoustic and/or subjective qualities (familiarity, frequency of exposure, etc.) of the sounds used could be undertaken. I will discuss this further with regard to the Experiment 3 results.

The greatest relative depth was between sounds that fell into different conceptual categories. Sounds in the same conceptual category had the least relative depth between them. Again, the results did not seem to be systematic, but such an "opposite" result from what was expected for the two extreme ends of the spectrum of conceptual similarity might be worthy of further consideration. Was this the result of the sounds that were chosen and how they were chosen, or is there some reason to expect that conceptually different sounds might exhibit greater relative depth between them than conceptually similar sounds? Perhaps the explanations are purely acoustic, but Chipley

(2004) did not find sufficient differences to pick it up. If so, then category is simply irrelevant.

There are some methodological problems that will have to be worked out as well. It seems virtually impossible to completely match overall sound level and perceived-distance to sound level profiles for multiple and sufficiently disparate environmental sound stimuli. However, an important attempt was made in the present study by conducting Experiment 3.

Experiment 3 Revisited

The findings of Experiment 3 may have been the most useful of the series. Admittedly, they were not the end goal of the present investigation. However, there are some important issues that arose out of the results that deserve further consideration. First, such a technique of examining changes in perceived-distance to changes in sound level profiles ($\Delta d'/\Delta dBA$ functions) for different sounds could prove to be a useful step in other ecological studies of perceived auditory distance. It was quite interesting to find that despite careful control of sound level, the various types of sound used could result in such different $\Delta d'/\Delta dBA$ functions.

The least squares linear fit lines on the $\Delta d'/\Delta dBA$ functions for each sound in Appendix D and the linear slope, and intercept values for those functions (see Table 7), were not originally considered in the selection of sounds. However, it is interesting in revisiting these data to note that the R-squared values for the sounds ranged from .58 to .97 for linear fit, suggesting a couple of things. First, there was a wide range in the suitability of linear models, despite the expectations from a traditional conceptualization

of a relative sound level cue. This was also true for intercept and slope, with ranges of -18.55 to 308.2 and 4.02 to -.22 respectively. Secondly, certain kinds of sounds (those with R-squared values closest to 1) are very predictable in terms of their $\Delta d'/\Delta dBA$ functions. Thus, they could make good candidate sounds for ecological studies.

Table 7: Intercepts and slopes for the $\Delta d'/\Delta dBA$ functions of each of the sounds from Experiment 3.

Sound	Intercept	Slope
Cow	26.9	-0.32
Harp	22.5	-0.27
Donkey	35.2	-0.44
Baby Cry	43.8	-0.56
Coughing	34.1	-0.43
Paper Crumple	26.4	-0.35
Whistling (Lips)	27.9	-0.32
Sheep	32.2	-0.39
Frog	18.6	-0.22
Clearing Throat	32.5	-0.39
Sawing	36.9	-0.43
Knocking	34.6	-0.41
Crow	98.8	-1.24
Dog Bark	71.8	-0.93
Doorbell	54.4	-0.69
Rooster	208.8	-2.71
Windchime	218.9	-2.99
Owl	308.2	-4.02

Given the careful control of the sounds, it seems likely that such wide ranges in fit statistics may be attributed, in part, to differences in the types of sounds. “Type” could refer to conceptual category as has been defined before. However, other factors like familiarity and frequency-of-contact, or “commonality” may be involved. Familiarity refers to a person’s past experience of the perception of the distance of a sound. What is

meant by “frequency-of-contact”, or what one might call “commonality”, is how often one may come into contact with a sound. To illustrate the important difference between the two, consider speech vs. a particular animal sound – the sound that a sheep makes. Most of us are very familiar with speech. We are also familiar with the “bah” sound that a sheep makes. However, most of us hear speech every day, but rarely hear sheep. Since acoustic differences between the sounds, conceptual similarity of the sounds, and familiarity all must be considered as contributing factors to perceived relative distance between sounds, commonality may also be an important variable to consider. While it was only considered as a possible variable *post hoc*, and thus, not intentionally measured or manipulated, commonality may very well be a factor that future investigations want to consider. However, a second look at the data (the $\Delta d'/\Delta dBA$ functions of the sounds) do not indicate any strong systematic differences simply based on the hypothesized commonality of the sounds used.

In Conclusion

With regard to practical applications of the discussed research, 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 many practical applications. Predicting how people will hear sounds in their environments is crucial for certain situations, such as virtual environments (where veridicality is key), cockpit auditory displays (where operator survival could be at stake), to understanding more prosaic examples (like navigating as a pedestrian through a large city). More generally, any user interface involving the use of sound could benefit from additional knowledge about how people

hear sounds in their environments. Performance errors could, in part, be explained by such knowledge. The goal of presenting sounds in a realistic and intuitive way in any system depends on extensive knowledge regarding how people hear sounds. 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. As well, more expanded areas of interest might be illuminated that involve how other species hear in their respective environments - for example, how Carolina wrens judge the distances of sound sources based on the structure of the sound signal (Wiley, 1998). The incorporation of more ecological methods will prove useful in these related endeavors.

References

- Ashmead, D. H., Davis, D. L., & Northington, A. (1995). Contribution of listener's approaching motion to auditory distance perception. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 239–256.
- Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 250-267.
- Barneclutt, P., & Pfeffer, K. (1998). Auditory perception of relative distance of traffic sounds. *Current Psychology: Developmental, Learning, Personality, Social*, 17, 93-101.
- Bonebright, T. L. (1996). An investigation of data collection methods for auditory stimuli: Paired comparisons versus a computer sorting task. *Behavior Research Methods, Instruments, & Computers*, 28, 275-278.
- Bonebright, T. L. (2001). Perceptual structure of everyday sounds: a multidimensional scaling approach. *Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29-August 1, 2001*, pp. 73-78.
- Butler, W. D., Levy, E. T., & Neff, W. D. (1980). Apparent distance of sounds recorded in echoic and anechoic chambers. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 745-750.
- Brungart, D. S., & Simpson, B. D. (2007). Cocktail party listening in a dynamic multitalker environment. *Perception & Psychophysics*, 69, 79-91.

- Cabe, P. A., & Pittenger, J. B. (2000). Human sensitivity to acoustic information from vessel filling. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 313-324.
- Campbell, D. T. & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171-246). Chicago: Rand McNally.
- Carello, C., Anderson, K., & Kunkler-Peck, A. (1998). Perception of object length by sound. *Psychological Science*, 9, 211-214.
- Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and two ears. *Journal of the Acoustical Society of America*, 107, 2224-2235.
- Chipley, M. R. (2004). *Does auditory similarity affect the usefulness of cues to perceived relative distance?* Unpublished master's thesis, N.C. State University, Raleigh, NC.
- Coren, S., & Enns, J. T. (1993). Size contrast as function of conceptual similarity between test and inducers. *Perception & Psychophysics*, 54, 579-588.
- Coren, S., & Hakstian, A. R. (1992). The development and cross validation of a self-report inventory to assess pure-tone threshold hearing sensitivity. *Journal of Speech and Hearing Research*, 35, 921-928.
- Ericson, M. A., Brungart, D. S., & Simpson, B. D. (2003). Factors that influence intelligibility in multitalker speech displays. *The International Journal of Aviation Psychology*, 14 (3), 313-334.

- Fowler, C. A. (1986). An event approach to the study of speech perception from a direct realist perspective. *Journal of Phonetics*, 14, 3-28.
- Fowler, C. A., & Rosenblum, L. D., (1991). Perception of the phonetic gesture. In I. G. Mattingly and M. Studdert-Kennedy (Eds.), *Modularity and the motor theory* (pp. 33-59). Hillsdale, NJ: Lawrence Erlbaum.
- Freed, D. J. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive events. *Journal of the Acoustical Society of America*, 87, 311- 322.
- Garbin, C. P. (1988). Visual-haptic perceptual nonequivalence for shape information and its impact upon cross-modal performance. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 547-553.
- Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2, 167-177.
- Gaver, W.W. (1993). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology*, 5, 1-29.
- Gibson, J. J. (1986). *The ecological approach to visual perception*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (Original work published 1979).
- Glave, R. D. & Rietveld, A. C. M. (1975). Is the effort dependence of speech loudness explicable on the basis of acoustical cues? *Journal of the Acoustical Society of America*, 58, 875–879.
- Goldstein, E. B. (2002). *Sensation and perception* (6th edition). Pacific Grove, CA: Wadsworth Thomson.

- Gygi, B., Watson, C. S., & Kidd, G. R. (1997). Factors in the detection of environmental sounds. *Journal of the Acoustical Society of America*, 101, Pt. 2, 3149.
- Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America*, 155, 1252-1265.
- Lakato, S., McAdams, S., & Causse, R. (1997). The representation of auditory source characteristics: Simple geometric form. *Perception and Psychophysics*, 59, 1180-1190.
- Lieberman, A., & Whalen, D. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4, 187-196.
- 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., Philbeck, J.W., & Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics*, 60, 966-980.
- Loomis, J. M., Lipka, Y., Klatzky, R. L., & Golledge, R. G. (2002). Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 335-345.
- Marcell, M. M., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Experimental and Clinical Neuropsychology*, 22, 830-864.

- Marcell, M. M. (2006). Sound stimuli. *Home Webpage for Michael Marcell*.
<http://www.cofc.edu/~marcellm/confrontation%20sound%20naming/zipped.htm>.
- 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., Desaulniers, D. H., Kiefer, S. A., Amerson, T. L., Jr., & Mills, J. T. (1981). Perceived loudness and visually-determined auditory distance. *Perception*, 10, 531-543.
- Mershon, D. H., & King, L. E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics*, 18, 409-415.
- Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 38, 379-387.
- Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, 395, 123-124.
- Neuhoff, J. G. (2004). *Ecological Psychoacoustics*. San Diego, CA: Elsevier, pp. 4-12.
- Neuhoff, J. G., & Heckel, T. (2004). Sex differences in perceiving auditory “looming” produced by acoustic intensity change. *Proceedings of the 2004 International Conference on Auditory Display, Sydney, Australia, July 6- July 9, 2004*.
- Rosenblum, L. D. (2004). Perceiving articulatory events: Lessons for an ecological psychoacoustics. In J. G. Neuhoff (Ed) *Ecological Psychoacoustics* . San Diego, CA: Elsevier, pp. 219-248.

- Russell, M. K., & Schneider, A. L. (2006). Sound source perception in a two-dimensional setting: Comparison of action and nonaction-based response tasks. *Ecological Psychology*, 18, 223-237.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327- 352.
- Vanderveer, N. J. (1979). Ecological acoustics: Human perception of environmental sounds. *Dissertation Abstracts International*, 40, 4543B. (University Microfilms No. 8004002)
- Warren, R. M. (1981). Measurement of sensory intensity. *Behavioral and Brain Sciences*, 4, 175-223.
- Warren, W. H., & Verbrugge, R. R. (1983). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 704- 712.
- Wiley, R. H. (1998). Ranging reconsidered. *Behavioral Ecological Sociobiology*, 42, 143-146.

Appendices

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

Scoring guidelines for 120 sounds listed alphabetically by sound level.

Marcell et al. (2000)

Sound label	A correct response must include...
Accordion	The root “accordion” or “concertina”
Airplane	The root “plane” or “jet;” reference to a specific type of jet (e.g., “747”) also acceptable
Baby crying	Both the roots “baby” and “cry;” synonym(s) may also be used (e.g., “infant bawling”)
Bagpipes	The root “bagpipe” or “pipes”
Banjo	The root “banjo”
Basketball	Either the root “basketball” or the word “ball” paired with the action of dribbling or bouncing
Bicycle bell	Both the roots “bell” and “bicycle” or “bike”
Birds chirping	The root “bird;” reference to the chirping sound alone not acceptable
Blinds	The root “blinds” (e.g., “Venetian blinds,” “miniblinds,” “opening the blinds”)
Blowing nose	Both the roots “blow” and “nose”
Boat horn	The roots “boat” and “horn” or a synonym (e.g., “foghorn;” “horn on a ship”); reference to a specific type of large ship (e.g., “tugboat;” “cruise ship”) also acceptable
Bongos	The root “bongo” or “conga;” reference to the root “drum” not acceptable
Bowling	The root “bowling”
Brushing teeth	Both the roots “brush” (verb) and “teeth;” “toothbrush” not acceptable
Burp	The root “burp” or “belch”
Camera	Either the root “camera” or reference to the action of taking pictures
Can crush	Both the object (any type of can) and the crushing action (e.g., “smashing can”)
Can opening	Both the object (e.g., “Coke,” “can,” “drink”) and the action of opening it (e.g., “opening a can,” “drink opened”); reference to the specific type of can (e.g., “soda pop top”) also acceptable
Car crash	Both the object (“car” or a synonym like “automobile”) and the event (e.g., “car crashing,” “automobile accident”)
Car horn	The root “horn;” reference only to the source of the sound (e.g., “car”) not acceptable

Sound label	A correct response must include...
Cash register	Both the roots “cash” and “register”
Cat	The root “cat” or a synonym (e.g., “kitten”); reference to the moving sound alone not acceptable
Chewing	Either the specific action (“chewing,” “munching,” “crunching”) or reference to any crisp food such as “carrot,” celery,” or “chips;” reference to only the general action of eating not acceptable
Chickens	The root “chicken” or “hen;” “rooster” not acceptable
Child coughing	The root “cough;” reference to the actor alone (e.g. “child”) not acceptable
Church bells	The root “bell” or “chime”
Clapping	The root “clap” or “applaud”
Clearing throat	Both the roots “clear” and “throat”
Coin dropping	The root “coin,” a synonym like “change,” or a particular type of coin; reference to “money” or only the action of dropping not acceptable
Cork popping	Either the root “cork” or a description of a specific object or action involving a cork (e.g., “champagne,” “popgun,” “opening a wine bottle”); other actions or objects not involving a cork (e.g., “balloon bursting,” “bubble gum pop”), a vague response (e.g., “pop”) not acceptable
Cow	The root “cow;” reference to the mooing sound alone not acceptable
Crickets	The root “cricket” or “grasshopper,” reference to a “bird chirping” sound also acceptable (although reference to birds alone is not)
Crow	The root “crow” or “bird;” reference to the other birds (e.g., “eagle”) or to the sound alone (e.g., “caw,” “squawking”) not acceptable
Crumpling paper	Both the crumpling action and the paper or plastic recipient of the action (e.g., “balling up paper,” “wrapper crinkling”); imprecise verbs (e.g., “rustling,” “rattling”) or reference to the object alone (e.g., “candy wrapper”) not acceptable
Cuckoo clock	The root “clock;” “cuckoo” alone not acceptable
Cutting paper	Either the object (“scissors”) or the cutting action plus recipient (e.g., “cutting paper”); reference to the cutting action alone or to a non-cutting action (e.g., “tearing paper” or “ripping”) not acceptable
Dog barking	The root “dog;” reference to the barking sound alone not acceptable
Donkey	The root “donkey,” “mule,” “ass,” or “burro;” reference to the sound alone (e.g., “hee-haw,” “braying”) not acceptable
Doorbell	The root “doorbell”
Door closing	Both the object (“door”) and the closing action (e.g., “door slamming,” “shutting door”)

Sound label	A correct response must include...
Drill	The root “drill;” “electric saw” or “power saw” (but not “saw” or “chainsaw”) also acceptable; a generic response (e.g., “power tool”) or reference to a household appliance (e.g., “hair dryer,” “blender”) not acceptable
Dropping ice in glass	The root “ice;” reference to the dropping action or to the glass receptacle without mention of ice not acceptable; reference to “marbles” combined with either the dropping action or the receptacle is acceptable
Drums	The root “drum”
Duck	The root “duck” or “goose;” reference to the quacking sound alone not acceptable
Elephant	The root “elephant”
Explosion	The root “explode;” reference to a large, man-made explosive sound (e.g., “bomb blast”) or the source of the sound (e.g., “cannon”) also acceptable; reference to only a small sound source (e.g., “gunshot”) or a natural explosive sound (e.g., “thunder”) not acceptable
Firecrackers	The root “firecracker” or “firework;” reference to guns or the sound of shooting not acceptable
Flute	The root “flute” or “recorder”
Frog	The root “frog” or “toad”
Frying food	The root “fry;” a general description of the event (e.g., “something cooking”) or reference to the object alone (e.g., “bacon”) not acceptable
Gargling	The root “gargle” or “gurgle”
Glass breaking	Both the object (anything made of glass) and the action of breaking (e.g., “broken glass,” “breaking a window”); reference to an imprecise object (e.g., “something breaking”) not acceptable
Gong	The root “gong;” reference to only the sound itself (e.g., “bong”) not acceptable
Guitar	The root “guitar”
Gunshots	Either the root “gun” or “shot;” use of a synonym (e.g., “rifle,” “pistol”) or reference to a specific type of hand-held, single-fire gun (e.g., “22 caliber”) also acceptable.
Hammering	The root “hammer;” responses such as “hitting nails” or “door knock” not acceptable
Harmonica	The root “harmonica” or “mouth organ”
Harp	The root “harp;” “harpsichord”, however, is not acceptable
Helicopter	The root “helicopter” or a synonym (e.g., “whirlybird,” “chopper”)
Horse galloping	The root “horse;” reference to the galloping sound alone not acceptable

Sound label	A correct response must include...
Jackhammer	The root “jackhammer”
Knocking	The root “knock;” reference to the recipient to the action alone (“door”) not acceptable
Laughing	The root “laugh” or a synonym for hearty laughter (e.g., “cackling,” “guffaw”)
Lawn mower	Both the roots “lawn” and “mower;” reference to other lawn tools (e.g., “weed eater”) or the general activity of “cutting grass” not acceptable
Lion	The root “lion” or “tiger;” a general descriptor (e.g., “big cat”) or reference to the roaring sound alone not acceptable
Machine gun	The root “machine gun,” “tommy gun,” or “semiautomatic;” reference to the concepts of “gunfire” or “rapid fire” alone not acceptable
Monkey	The root “monkey,” “chimp,” or reference to another non-human primate; reference to the sound alone (e.g., “screech”) not acceptable
Mosquito	The root “bee,” “fly,” or “mosquito,” or reference to any buzzing, flying insect; reference to the buzzing sound alone, or to the generic “bug” or “insect,” is not acceptable
Motorcycle	The root “motorcycle” or a synonym (e.g., “chopper”)
Ocean	The root “ocean,” “wave,” or a synonym (e.g., “sea”); “beach,” without reference to water, is not acceptable
Owl	The root “owl”
Piano	The root “piano” or “keyboard”
Pig	The root “pig”
Pinball	The root “pinball;” reference to “game” or “machine” alone not acceptable
Ping-Pong	The root “Ping-Pong” or “table tennis” (but not “tennis” alone)
Police siren	Either the root “siren” or reference to an emergency vehicle that uses a siren (e.g., “police car,” “ambulance”); reference to “police” alone not acceptable.
Pouring water	At least two of the following concepts; the action (“pouring” or “filling”), the liquid being poured (any liquid), and the receptacle (any container) (e.g., “filling cup,” “pouring liquid”)
Pullchain lightswitch	Either the root “pullchain,” “light,” or “switch;” reference to “stapler” or “pen clicking” also acceptable
Rain	The root “rain;” “shower” not acceptable due to potential confusion with bath shower
Rattlesnake	Either the root “snake” or “rattler;” “rattle” alone not acceptable

Sound label	A correct response must include...
River	The root “river,” “stream,” “brook,” “creek,” or reference to any natural, flowing body of water, reference to “waterfall,” or an imprecise response like “running water” not acceptable
Rooster	The root “rooster” or “cock;” reference to the sound alone (e.g., “cock-a-doodle-doo”) not acceptable
Sandpaper	The root “sandpaper” or “sanding”
Sawing	The root “saw”
Saxophone	The root “sax”
Scream	The root “scream”
Seal	The root “seal” or “sea lion;” reference to the barking sound alone not acceptable
Sheep	The root “sheep,” “lamb,” or “goat;” reference to the baaing sound alone not acceptable
Shuffling cards	Either the root “shuffle” or the root “card”
Sneeze	The root “sneeze”
Snoring	The root “snore;” reference to just the act of sleeping not acceptable
Sonar	The root “sonar,” “submarine,” or “radar”
Stapler	The root “stapler”
Swords	Either the root “sword,” a synonym (e.g., “saber”), reference to a large, sharp cutting tool like a machete, or the action of sharpening a knife; reference to the sound alone (e.g., “clanging metal”) or to “knife” or “knives” either alone or combined with a word other than “sharpening” (e.g., “knives scraping”) not acceptable
Tea kettle	The root “kettle” or a synonym (e.g., “teapot”); “boiling water,” a description of the whistling sound alone, or an imprecise reference to the object (e.g., “steam from a pot”) not acceptable
Tearing paper	Both the object (any paper or cloth object) and the action of tearing it (e.g., “cloth ripping,” “paper tear”); an imprecise reference to an object (e.g., “something tearing”) not acceptable
Telephone	The root “phone;” reference to the ringing sound alone not acceptable
Thunder	The root “phone;” reference to the ringing sound alone not acceptable
Toilet flushing	The root “toilet” or a synonym (e.g., “commode”); reference to the flushing sound alone not acceptable
Train	The root “truck” or “bus;” reference to a particular type of large truck (e.g., “18-wheeler”) acceptable, but “bulldozer” or other construction machinery is not; imprecise responses like “car” or “vacuum cleaner” not acceptable

Sound label	A correct response must include...
Trumpet	The root “trumpet” or any trumpet-like brass wind instrument (e.g., “bugle”); use of the generic label “horn” also acceptable
Turning pages	Both a turning action and a paper object (e.g., “changing pages,” “flipping papers”); imprecise reference to the action (e.g., “rustling papers”) not acceptable
Typewriter (manual)	The root “typewriter;” reference to “cash register” not acceptable
Velcro	The root “Velcro” or reference to the concept of tearing cloth (e.g., “fabric ripping”)
Violin	The root “violin,” a synonym (e.g., “fiddle), or the root “viola;” “string instrument” or “cello” not acceptable
Water bubbling	The root “bubble” or “boil;” reference to a circumstance in which air is blown into a liquid (e.g., “fish tank,” “blow through straw”) also acceptable; reference to the liquid alone not acceptable
Water draining	Either the object (any liquid) or vehicle (“sink,” “tub”) and the act of draining (e.g., “water going down the drain,” “draining a liquid,” “tub emptying”); reference to other specific actions with water (e.g., “flushing,” “pouring water,” “running water”) not acceptable
Water dripping	The root “drip” or “drop;” reference to the liquid alone (e.g., “water”) or to rain not acceptable
Whip	The root “whip;” reference to a gunshot sound or description of the crackling sound alone not acceptable
Whistle (instrument)	The root “whistle”
Whistle (lips)	The root “whistle;” reference to an event without the root word (e.g., “calling a dog”) not acceptable
Wind	The root “wind” or any synonym for a strong wind (e.g., “gale”)
Wind chimes	Both the roots “wind” and “chime;” “bells,” “windbells,” or “chimes” alone not acceptable
Wolf	The root word “wolf” (e.g., “werewolf”) or “coyote;” reference to “wild dogs” or the howling sound itself not acceptable
Woodpecker	The root “woodpecker”
Yawning	The root “yawn”
Zipper	The root “zip”

Appendix C

Perceived relative distances for participants exposed to reference sounds 1 and 2 (for each of 15 conditions), and verbal judgments of the egocentric distance of the reference sound in m.

Men (Reference 1)

Condition	Sound Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean	SD
Ref. 1 Lower	54	160	200	200	200	225	200	300	120	145	200	150	180	175	175	150	185	42
Ref. 1 Same	60	110	100	100	100	100	100	100	100	95	80	100	95	100	100	125	100	9
Ref.1 Higher	66	30	40	50	25	30	25	15	50	60	50	50	50	40	50	10	38	15
Same Lower	54	110	150	50	150	115	120	150	120	115	150	175	90	220	100	200	134	43
Same Same	60	80	50	20	25	80	75	75	50	80	80	80	80	80	65	250	78	52
Same Higher	66	30	10	1	5	45	20	25	20	15	50	25	40	90	50	100	35	29
Similar Lower	54	140	150	110	100	130	170	150	120	140	150	125	120	210	225	25	138	46
Similar Same	60	85	100	70	75	75	100	100	110	100	100	120	110	90	100	75	94	15
Similar Higher	66	50	60	30	50	50	70	50	20	80	80	100	50	67	50	50	57	20
Different 1 Lower	54	140	125	150	200	175	150	300	80	175	200	175	135	240	225	5	165	69
Different 1 Same	60	75	100	90	100	115	75	100	50	110	200	115	90	120	100	25	98	38
Different 1 Higher	66	30	30	25	25	40	70	50	20	35	100	85	90	60	75	10	50	28
Different 2 Lower	54	150	200	150	200	210	125	300	30	160	50	200	120	200	200	10	154	77
Different 2 Same	60	50	50	75	50	55	50	50	20	55	120	90	100	70	50	80	64	25
Different 2 Higher	66	30	60	10	50	65	50	75	10	20	75	20	78	75	25	70	48	26

**D' of Reference
(converted to m)**

3.05 3.66 7.62 1.52 3.66 3.05 2.74 3.05 0.91 3.66 6.10 4.57 2.59 3.05 4.57 3.59 3.62

Men (Reference 2)

Condition	Sound Level	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Mean	SD
Ref. 2 Lower	54	200	300	175	200	200	175	300	125	200	145	250	160	200	300	175	207	56
Ref. 2 Same	60	100	100	125	100	100	100	100	100	100	100	100	100	100	100	100	102	6
Ref. 2 Higher	66	50	25	25	50	17	75	25	50	50	50	10	10	40	50	50	38	19
Same Lower	54	175	250	350	200	160	200	300	200	150	175	250	200	80	200	200	206	64
Same Same	60	100	200	150	200	125	75	150	185	50	120	125	100	80	125	125	127	44
Same Higher	66	50	25	50	100	85	60	50	90	25	25	80	10	10	50	50	51	29
Similar Lower	54	175	250	250	250	164	150	200	250	250	175	175	160	180	200	200	202	38
Similar Same	60	100	150	125	150	100	100	100	100	175	125	120	100	130	110	120	120	23
Similar Higher	66	25	25	50	100	20	75	90	85	50	30	100	60	100	70	80	64	29
Different 1 Lower	54	200	350	200	250	220	200	300	200	275	150	200	160	180	275	225	226	55
Different 1 Same	60	100	200	100	150	90	125	130	135	125	115	80	110	150	150	100	124	31
Different 1 Higher	66	25	100	50	95	64	75	75	75	100	30	25	40	60	60	75	63	25
Different 2 Lower	54	200	300	300	250	200	200	200	225	250	200	160	140	130	250	200	214	50
Different 2 Same	60	50	150	100	100	130	125	100	95	50	125	75	50	130	80	100	97	32
Different 2 Higher	66	25	50	50	100	70	100	50	65	25	80	25	10	50	75	50	55	27

**D' of Reference
(converted to m)**

1.52	1.52	1.22	1.37	0.91	1.83	2.13	3.05	1.83	1.83	4.57	4.57	1.83	0.91	1.52	2.04	1.15
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Women (Reference 1)

Condition	Sound Level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean	SD
Ref. 1 Lower	54	300	150	200	250	200	150	210	200	200	200	200	250	180	130	200	201	43
Ref. 1 Same	60	100	100	100	100	100	100	110	100	100	100	100	100	100	100	100	101	3
Ref. 1 Higher	66	80	15	10	70	30	10	50	50	50	50	5	25	50	20	50	38	23
Same Lower	54	300	175	125	110	110	60	75	200	100	100	10	125	100	200	100	126	69
Same Same	60	50	50	20	20	55	30	75	100	125	50	50	50	75	30	50	55	29
Same Higher	66	200	15	20	10	25	10	15	50	25	50	5	10	10	15	25	32	48
Similar Lower	54	200	125	125	125	200	85	225	200	200	150	250	200	137	80	150	163	51
Similar Same	60	100	75	150	180	100	85	75	125	150	50	100	80	150	50	100	105	39
Similar Higher	66	80	175	30	50	50	75	75	50	75	25	25	50	100	20	100	65	40
Different 1 Lower	54	400	125	150	200	200	80	200	200	150	200	200	200	250	200	150	194	70
Different 1 Same	60	100	110	100	200	60	50	125	150	100	20	125	100	125	50	120	102	45
Different 1 Higher	66	90	87	10	50	40	20	125	75	50	100	40	15	75	80	75	62	33
Different 2 Lower	54	150	175	200	10	250	120	175	200	200	150	120	200	200	150	175	165	55
Different 2 Same	60	100	90	75	30	50	50	10	100	75	150	90	75	50	30	60	69	35
Different 2 Higher	66	50	10	10	120	50	10	10	75	75	50	10	30	25	20	50	40	32

**D' of Reference
(converted to m)**

4.57	1.07	2.74	7.62	3.66	0.61	7.62	3.66	1.22	2.13	3.05	2.44	3.05	6.10	1.37	3.39	2.23
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Women (Reference 2)

Condition	Sound Level	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	Mean	SD
Ref. 2 Lower	54	200	200	200	130	300	135	200	250	200	200	200	200	200	200	200	201	39
Ref. 2 Same	60	100	100	100	100	100	100	100	100	100	125	100	100	100	100	100	102	6
Ref. 2 Higher	66	50	5	25	40	50	75	50	25	50	50	25	50	25	50	40	41	17
Same Lower	54	300	200	200	150	300	80	175	300	200	200	150	200	200	175	200	202	60
Same Same	60	200	100	100	120	90	110	100	150	100	125	70	115	100	100	110	113	30
Same Higher	66	49	10	50	105	10	40	50	50	50	50	25	80	50	50	75	50	25
Similar Lower	54	400	150	200	200	300	160	175	200	100	150	180	150	150	175	150	189	73
Similar Same	60	200	100	150	110	90	90	100	150	100	150	70	100	200	100	100	121	40
Similar Higher	66	100	75	75	100	60	60	25	50	50	100	60	50	50	25	50	62	24
Different 1 Lower	54	325	200	250	160	300	120	200	200	200	200	180	150	200	200	200	206	52
Different 1 Same	60	300	125	150	125	200	120	100	200	75	125	150	125	125	100	125	143	55
Different 1 Higher	66	50	25	50	90	50	80	25	50	50	75	90	40	50	25	50	53	22
Different 2 Lower	54	400	100	200	140	250	200	200	300	200	200	200	175	150	200	200	208	70
Different 2 Same	60	100	50	100	60	70	120	50	100	75	100	80	110	100	50	100	84	24
Different 2 Higher	66	50	5	50	95	20	30	25	75	25	25	110	50	25	25	25	42	30

**D' of Reference
(converted to m)**

0.30	3.05	1.52	3.05	6.10	3.05	1.52	1.52	9.14	1.22	4.00	3.66	3.05	1.52	2.44	3.01	2.21
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Appendix D

Individual graphs of median perceived distances for each sound at different sound levels.

