Downloaded from orbit.dtu.dk on: Dec 17, 2017

Technical University of Denmark

Information from Impact Sounds: Normal and Impaired Hearing

Kirkwood, Brent Christopher; Poulsen, Torben; Naylor, Graham

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Kirkwood, B. C., Poulsen, T., & Naylor, G. (2007). Information from Impact Sounds: Normal and Impaired Hearing.

DTU Library
Technical Information Center of Denmark

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Information from Impact Sounds: Normal and Impaired Hearing

Ph.D. thesis by

Brent C. Kirkwood

2007

Technical University of Denmark

Contents

Preface

This publication was submitted as partial and final fulfillment of the requirements for obtaining the degree of Doctor of Philosophy (Ph.D.) in Electronics and Communications at the Technical University of Denmark. The work described within was completed in the Acoustic Technology section of the Ørsted·DTU department at the Technical University of Denmark between October 1, 2003 and October 29, 2006. A public defense was held on March 9, 2007. The project was supervised by Associate Professor Torben Poulsen of the Technical University of Denmark and co-supervised by Graham Naylor, Director of Research for Oticon A/S. The project was funded by a scholarship from the university. Ethical considerations for conducting listening experiments in the project were made with reference to the Scientific Ethics Committee of Copenhagen County file number KA 04159g, as granted to the Centre for Applied Hearing Research of the Technical University of Denmark.

I have attempted to write this dissertation in such a way that it can be understood by people familiar with classical approaches to the study of human hearing. Although I have not expected them to be my primary audience, I hope that it is also accessible to those with knowledge of human perception in general, particularly ecologically based approaches to perceptual research.

This document would not be in the state that it is without the influence of many people. I would like to thank my supervisors, Torben Poulsen and Graham Naylor, for being available for discussions during the project. For originally suggesting the topic of the project, of which I had no prior knowledge, I am thankful to my co-supervisor Graham Naylor. I thank Torsten Dau of the Centre for Applied Hearing Research at the Technical University of Denmark for his encouragement and willingness to collaborate during the project. Even if I do not mention all of them by name, I am grateful to many of my colleagues in the section of Acoustic Technology, who have been of tremendous help to me. Special thanks go to Iris Arweiler and Erik Schmidt for their audiological advice. Thanks go to those who proofread and provided suggestions concerning parts of this dissertation including

Erik Schmidt, Adam Weisser, and Torsten Dau. I also wish to thank the test subjects who participated in my listening tests. Without their ears, there would be only speculation.

Thanks also go to Oticon, the company, for providing hearing aids for my listening tests, and to Elisabet Sundewall Thorén of Oticon for her assistance in programming the hearing aids. I very much appreciate the invaluable encouragement, references, and advice of Alan Costall, Professor at the University of Portsmouth. I would also like to thank Daniel Russell, Associate Professor at Kettering University, for originally sparking my interest in the scientific study of acoustics. Without his enthusiasm, I would have likely never been interested in the subject. Finally, I should thank all of those who have done auditory research before me. Whether the ideas in this thesis are presented as my own or cited as the work of others, my thoughts have no doubt been influenced by the work of many people.

Among other reasons, I have enjoyed this project because many of the ideas in it are a bit provocative. I hope that readers are pushed to contemplate topics that they have not yet considered.

Brent C. Kirkwood April 26, 2007

Summary

In their everyday lives, people gather an abundance of information from the sounds present in the environments in which they act. With the exception of speech, these sounds have generally not been considered by psychoacousticians as providing information to listeners, but as producing sensations such as loudness, pitch, and timbre. It is first argued that this oversight must be addressed in order to provide an efficient and accurate means of understanding the real-world functionality of the human auditory system. Impact sounds were then chosen for this project as a common everyday sound for experimental investigation.

In the first experiment, the influence of stimulus presentation method was investigated. Using only acoustic cues, normal-hearing subjects estimated the lengths of wooden rods dropped onto a linoleum floor. Their length estimation accuracy was compared for three presentation methods: 1) live presentation, 2) headphone playback of binaurally recorded stimuli, and 3) headphone playback of monophonically recorded stimuli. Subjects made larger errors when listening to monophonically recorded stimuli than when performing the task live. Binaural recordings were not found to produce results that were worse than live presentation. The results indicate that spatial cues may be an important aspect in auditory length perception, and that the selection of an inappropriate presentation method can deprive listeners of information normally available to them in the real world.

The second experiment was an investigation of whether hearing-impaired listeners are as capable as normal-hearing listeners in hearing three ecologically relevant properties of impact sounds resulting from dropped rods: 1) the materials of the rods, 2) the lengths of the rods, and 3) the heights from which the rods are dropped. The results of listening tests are presented in which two subject groups, normal hearing and hearing impaired, have been tested with and without hearing aids. Hearing-impaired subjects without their hearing aids were found to perform worse, as a group, at judging the three parameters. Equipped with hearing aids, they remained worse than the unaided normal-hearing subjects at judging only material. The results are therefore informative about the abilities of normal-hearing and hearing-impaired listeners, and about the influence of the hearing aids.

Finally, an experiment in auditory scene analysis was conducted in an attempt to help bridge the gap between classical psychoacoustics and an information-based approach to auditory perception. Unfortunately, the results were inconclusive and therefore not promising for bridging this gap. However, the outcome of this experiment emphasizes the need to use an ecological approach for studying the human auditory system.

Danish Summary

Titel: Informationer fra slaglyde: Normal og nedsat hørelse

I hverdagen modtager folk en overflod af informationer fra lyde i deres omgivelser. Med undtagelse af tale har psykoakustikere generelt ikke beskæftiget sig med, at disse lyde forsyner lytterne med informationer, men med sanseindtryk såsom hørestyrke, tonehøjde og klangfarve. Der argumenteres for, at dette bør tages i betragtning for at forstå hørelsens funktion i den virkelige verden. I dette projekt er det valgt at foretage eksperimentelle undersøgelser af slaglyde som en almindelig hverdagslyd.

I det første forsøg blev indflydelsen af stimuluspræsentationsmetoden undersøgt. Ved kun at anvende akustiske informationer bedømte normalthørende forsøgspersoner længden af træpinde, der faldt ned på et linoleumsgulv. Længdevurderingen blev sammenlignet ved tre præsentationsmetoder: 1) Live præsentation 2) hovedtelefon-gengivelse af binauralt optagede stimuli og 3) hovedtelefon-gengivelse af monooptagede stimuli. Testpersonerne lavede flere fejl, når de lyttede til mono optagede stimuli, end når de lyttede live. Resultaterne af de binaurale optagelser var ikke værre end live præsentationen. Resultaterne indikerer, at rumlige informationer kan være et vigtigt aspekt i auditiv længde-perception, og at valget af en uhensigtsmæssig præsentationsmetode kan berøve lyttere informationer, som normalt er tilgængelige for dem i den virkelige verden.

I det andet forsøg blev det undersøgt, om hørehæmmede lyttere er lige så kapable som normalthørende til at høre tre økologisk relevante egenskaber fra stænger, der var faldet ned: 1) stængernes materiale 2) stængernes længde og 3) hvilken højde de var faldet ned fra. Resultater af lytteforsøg er præsenteret, hvor to grupper af forsøgspersoner, normalthørende og hørehæmmede, er blevet testet med og uden høreapparater. Hørehæmmede forsøgspersoner uden deres høreapparater klarede sig dårligere som gruppe, når de skulle bedømme de tre parametre. Udstyret med høreapparater forblev de dårligere end de normalthørende forsøgspersoner uden høreapparat, når de skulle bedømme materialet. Resultaterne er derfor informative, hvad angår normalthørendes og hørehæmmedes evner og høreapparaters betydning.

Afslutningsvis blev der udført et eksperiment i auditory scene analysis som et forsøg på at bygge bro mellem klassisk psykoakustik og informationsbaseret auditiv perception. Desværre var resultaterne ufyldestgørende og derfor ikke lovende med hensyn til at bygge denne bro. Imidlertid understreger disse resultater nødvendigheden af at anvende en økologisk tilgang for at undersøge det menneskelige auditive system.

Chapter 1

Introduction

> *Reality is always richer than the imagination.*
>
> -William Mace (2005)

For one reason or another, most people tend to think of events in the world as visual (Jenkins, 1985). "Seeing is believing," or so the saying goes (Neuhoff, 2004d). The sense of hearing is acknowledged to provide sensory input, but there is a tendency for it to be treated as a secondary provider of information. In actuality, events in the world commonly provide information for multiple perceptual systems, simultaneously. They structure light, sound, chemical properties of their surrounding environment, vibrations, and more (Jenkins, 1985).

The human perceptual systems that are capable of interpreting this information are at the same time redundant and independent. They work together to take advantage of the world's information in its various formats. The relatively short wavelengths of light, for example, make fine spatial resolution possible, while the longer wavelengths of sound make hearing around corners possible (Griffin, 1959; Neuhoff, 2004a). The long wavelengths of sound help make it possible to gather information from all directions without reorientation. Just as the visual system can provide information that may not be available to the auditory system, acoustic signals may convey information about material, structure, or other properties of objects that may not necessarily be readily apparent to the visual system (McAdams, Chaignec & Roussarie, 2004). Furthermore, in conditions where information is available to multiple sensory systems, perception may be enhanced above that of either sense alone (Neuhoff, 2004d).

Before attempting to understand *how* animals recognize sounds, it makes sense to first examine *what* those animals use sounds for (Heine & Guski, 1991). It makes little sense to attempt to reach a deep understanding of the physiology of an auditory system before considering what kind of acoustic environment it is exposed to and to what information it collects from its environment. Similarly, modeling systems that are not yet understood seems premature (Bregman, 1990). Producing models of systems whose purpose has not yet been fully explored is bound be inefficient, resulting in models of phenomena that do not actually happen outside of laboratories. It is a noble goal to have perfect mathematical models of the human auditory system, but there is much work to be done to first understand for what the system is used. Prior knowledge of *what* an auditory system is used for can be used to guide an investigation of *how* the perceptual system performs its basic tasks.

There is a great deal of information constantly being presented to the human auditory system, but historically, research on the human auditory system has primarily been focused on sensation and not the perception of information. The fact that sounds provide information to listeners has typically been ignored (Gibson, 1966). It has simply been assumed that sensation is directly linked to perception. A microscopic, atomistic approach to studying hearing has dominated the projects of laboratories around the world. Hopes have been that an understanding of every possible minute phenomenon will lead to an overall understanding of how the auditory system uses acoustic information in the real world, but it is not clear if this is even possible (Bregman, 1990). Such an understanding of the auditory system – the way it works, not just observations of the input and output – would of course be tremendously valuable, but this understanding has not been reached (Plomp, 2002). It may therefore be wise to approach the problem from another angle.

It has been suggested by Plomp (2002) and others (e.g., VanDerveer, 1979) that the reason hearing research has been focused on sensation is because of historical and technological developments. The Fourier theorem, both synthesis and analysis, has had a profound influence on auditory science. Textbooks on hearing typically explain that all sounds can be considered to be built up by Fourier synthesis (Noble, 1983). The idea of the ear as a frequency analyzer has been generally accepted since von Helmholtz (1954, 1863) applied it to his work (Plomp, 2002). Since this time, there has been an overwhelming dominance of sinusoidal tones in hearing research. Other technological constraints may have partly controlled the direction of hearing research. Signal generation, signal recording, and playback techniques have certainly influenced experimental methods of hearing researchers. For example, until the recording and presentation of sound became simple, there were limits to the kinds of sounds that could be generated artificially. As another example, it may be no coincidence that most research concerning auditory localization has occurred with a focus on the horizontal plane. This plane happens to be easy to simulate with headphones (Neuhoff, 2001a). Of course, it has always been possible to produce live stimuli, but the idea of using such realistic sounds seems to be a novelty in hearing research.

Tools such as signal generators, headphones, anechoic chambers, and soundproof booths along with highly developed experimental methods and analysis techniques have encouraged the experimental study of hearing (Neuhoff, 2004c). For the most part, these tools have been and will continue to be of great benefit. The tools have enabled researchers to verify the sensations reported by listeners. Such results can be trusted. They describe perfectly valid phenomena. A high level of understanding has been reached concerning these types of phenomena – much higher than that of the perception of real-world sounds – but it appears that these facts have been assumed to provide the basis for perception (Gibson, 1966).

The austerity and exactness that have been promoted through classical experiments have made it seem that natural hearing cannot be studied, that the classical approach is the only way to study hearing. The inability to control and isolate all variables in a natural environment is apparently a reason that prevents experiments from being conducted under natural conditions, but whether or not this rigidity is always necessary, possible, or even beneficial is debatable (Walker & Kramer, 2004). Are these constraints, such as isolated variables, found in the real world? Should static, acoustically impoverished signals with no intrinsic meaning, be the focus of attempts to understand auditory perception? It is more likely that a considerable amount of research is first required before stimuli that are worth testing can be artificially created (VanDerveer, 1979).

Much has been learned about the physiology of the ear, about sensations of loudness and pitch in response to mathematically simple stimuli, about the thresholds required to provoke these sensations, about the ear's temporal resolution, and about its frequency selectivity. These findings are no doubt useful in understanding the function of the ear as an organ, but few of them have helped to explain real-world listening. Gibson (1966) provided one of many warnings, 40 years ago, though the words are still acutely relevant today:

> When the senses are considered as channels of sensation (and this is how the physiologist, the psychologist, and the philosopher have considered them), one is thinking of the passive receptors and the energies that stimulate them, the sensitive elements in the eyes, ears, nose, mouth, and skin. The experimenters in physiology and psychology have been establishing the conditions and limits at this level of stimulation for more than a century. A vast literature of sensory physiology has

> developed and a great deal is known about the receptors. It is a highly respected branch of science. But all this exact knowledge of sensation is vaguely unsatisfactory since it does not explain how animals and men accomplish sense perception.

An argument for using the mathematically simple stimuli that are typically used in hearing research is that the results from more complicated signals would be impossible to understand. Energy from mathematically simple stimuli is no doubt the easiest kind of energy to measure and characterize, but energy without structure lacks information. Energy structured in mathematically complex ways seems to be that most rich in information (Gibson, 1966). Similarly, an argument for studying small, manageable "subsystems" of the auditory system has been that anything more would be too complex. While the truth of such claims is questionable, there can be no hard feelings towards researchers who desire to stick to what they feel can be accomplished.

However, there is no reason to believe that the human auditory system has more trouble dealing with mathematically complex signals than mathematically primitive signals. The very opposite may be true. As an example, it has been noted that human auditory localization[1] is better when attempting to localize "complex" sounds (e.g., wide-band noise) than "simple" sounds (e.g., sinusoids). As put by Jenkins (1985), "There may be an important lesson here." Natural signals, of potentially great mathematical complexity, may be the easiest to understand for the human auditory system – a system that has evolved to resonate to that information, not to synthetic pure tones presented in laboratory environments (Johansson, 1985; Jenkins, 1985; Plomp, 2002). To reiterate, that which is complex for physics may be simple for perception (Gibson, 1966).

Research on temporal segregation of auditory streams is claimed to be an attempt to bridge the gap between classical psychoacoustics and hearing in complex, real-world environments (e.g., Deutsch, 1982; Bregman, 1990; Cusack & Carlyon, 2004). Such a goal is certainly to be applauded. However, the stimuli continuing to be used in this research are far from being like those in the real world. Typical stimuli tested, such as pure tones of rapidly alternating frequency, seem to have little basis in reality. Research involving other ways of more realistic temporal modulation of signals is an encouraging step, but the tasks required of the test subjects in these experiments provide results that say little about listening outside of the experiments. It is doubtful that such listening conditions frequently exist outside of laboratories. As put clearly by Gaver (1988), "The aim of such research

---

[1] The ability to localize sounds is one real-world use of the human auditory system that has been somewhat explored.

seems to be to find how the traditional attributes[2] of sound can be manipulated to give rise to stream segregation." None of this is to say that a desire to understand the abilities of humans to separate sound sources is not needed. On the contrary, this ability of humans is such a basic skill[3] that it should be considered from the beginning (Plomp, 2002). The results of traditional streaming research may one day be useful to understanding the perception of everyday sounds. It is, however, unlikely that such a use will be found without tests of simultaneous *ecologically valid* sounds presented in more real-world-like sound fields (Gaver, 1988). Other auditory scene analysis topics such as segregation of simultaneous inputs may provide results that are more useful. The conditions typically tested in experiments of this nature are more realistic than those of temporal streaming research; however, they too need constraints imposed by real-world sounds.

Speech is one example of a real-world auditory stimulus that *has* been studied. Though much of the research has been in vain (Plomp, 2002), valuable lessons may be available that can be used in future research of both everyday sounds and speech (Rosenblum, 2004). Jenkins (1985), Plomp (2002), and Rosenblum (2004) have each advocated that lessons for the study of everyday sounds can be learned from the study of speech. While speech perception is far from being completely understood, its research history is much longer, and much more ecologically valid than that of any other type of sound. Both the successes and the failures of speech research may help guide the study of everyday sounds.

Jenkins (1985) points out two related points from speech investigations that may suggest that auditory perception is worth approaching from multiple angles. First, he reminds the reader that, "Speech signals that are radically different in terms of the usual physical analysis may be heard as 'the same thing'." As if that point alone is not reason enough to indicate that a different research paradigm might be useful, he continues by stating, "Speech signals that are remarkably similar in terms of the usual physical analysis may be heard as 'different things'." It has also been shown that the best predictor for the perceived loudness of speech is not any simple acoustic dimension or combination of dimensions, but a property of the source itself: articulatory effort (Rosenblum & Fowler, 1991). It is clear from these points that there are shortcomings in the understanding of perceptually-relevant signal attributes.

Apart from speech, music is an additional popular topic of hearing research, but for classification purposes, it can be considered as a sort of perceptual luxury and not a signal of importance to survival (Gibson, 1966). It is therefore curious that more focus has not been placed on everyday sounds, which are certainly more common

---

[2] Traditional attributes of sound include loudness and pitch, for example.

[3] As is directional hearing.

in a person's environment than either speech or music. The terms "everyday sounds" or "environmental sounds" can be used synonymously to describe all of the sounds present in the normal everyday environments of listeners, whether or not the sounds are man made, but specifically excluding speech and music. The category of everyday sounds encompasses a wide variety of sounds that provide a great deal of information to listeners in their daily lives. Everyday sounds are constantly occurring, and while speech is frequent, there are (fortunately) pauses. It may be argued that verbal communication with others of the species is of greater importance. This fact is debatable, but it is certain that everyday sounds, the broad category of non-speech and non-music sounds, have been neglected by the auditory research community (Carello, Wagman & Turvey, 2005).

What can explain the lack of interest in the perception of everyday sounds? Are they so common and obvious that their utility is taken for granted (Jenkins, 1985)? Are they so easy to recognize that they are not worth studying? Is the category of "everyday sounds" so large that tackling it is too daunting? Are these types of sounds of such great variety that people have difficulty in discussing them? Are they of little importance? Are they simply background noise making concentration and the perception of other signals of interest more difficult? Are they simply believed to be too complicated to possibly understand? VanDerveer (1979) suggested that because they are often correctly perceived, there may be nothing curious about them that would provoke imaginative theorizing. Possibly many of these issues are partly responsible for the neglected topic of everyday sound research. This project has been an attempt to help address this deficiency.

The focus of this thesis is on impact sounds. Impact sounds were chosen as one class of everyday sounds because of their prevalence in the world and because they are easy to generate. Impact sounds make up a large quantity of the sounds present in a person's environment. Whenever two objects, organisms, or combination thereof collide, the objects and organisms deform, vibrate and produce vibrations in the air surrounding them. These vibrations, a potential stimulus, may be called "sound" if there are listeners in the area who are capable of detecting them (Gibson, 1966). Soft collisions may be heard in the form of very quiet impact sounds, for example from a person typing, a person setting a pen on a desk, from an insect flying into a window, or from even quieter impacts. Other types of impact events, such as a person setting a glass down on a table, a person walking upstairs, a book cover being closed, a carpenter hammering a nail into a wall, a car door being closed by a guest who has just arrived in a listener's driveway, and others, may be at a variety of levels depending on the force of the impact, proximity, and other variables. Even louder impacts such as a vase falling off a shelf or a child falling over in his or her chair may alert a listener that sudden action may be necessary. When audible, all of these sounds provide information to people capable of detecting it. This thesis will focus on the types of information and the accuracy of the information provided to people by such sounds.

For the purposes of this project, the ecological approach to perception, advocated by James Gibson (1966, 1979), has provided a rough framework to guide questions and to guide the ways of answering these questions. Because the words "ecology" and "ecological" can have multiple meanings and implications, a short explanation of the use of these words is in order. For this thesis, the word "ecological" will be used to imply a consideration for the interaction of an organism and its environment. To be clear, the word "ecological", or more generally "ecology", will not be used to describe a philosophy of "acoustic ecology" in which the preservation of acoustically pleasant, natural sounds are advocated (e.g., Wrightson, 2000; Kettles, 2006). The word "ecology" or derivatives will not imply a desire to protect nature or reduce man's influence on nature. "Ecology" will simply refer to research concerning the interaction of an organism (e.g., a person) and its environment. "Ecological acoustics" or an "ecological approach" to auditory perception will refer to a way of studying a perceptual system that considers the environment as an important part of the way in which an organism acts in daily life.

Although the ideas of ecological psychology as proposed by Gibson (1966, 1979) have inspired much of the work in this thesis, further use of the word "ecological" in this thesis will not imply a strict adherence to the ecological psychology approach described by James Gibson (1966), nor that of Egon Brunswik (for a comparison of the approaches, see Vicente, 2003). Nevertheless, many of the principles appear to serve as an excellent guide for an alternative approach to studying hearing, an approach which considers that listeners are living, perceiving, and acting beings in a world, and that these listeners use their hearing to guide and perform these tasks.

1.1 Applications

Knowledge of how humans perceive impact sounds and everyday sounds in general can be used in many ways. Some of the most direct applications are in auditory icons, assistive listening devices, auditory displays, machine "hearing", sound design, video games, and other virtual reality simulations.

Auditory icons, which may be used by computers or in other man-machine interfaces, can be utilized to efficiently provide information to users in the form of confirmation and other feedback concerning the status of the task that they are performing (e.g., Gaver, 1988). As opposed to "earcons" (e.g., Blattner, Sumikawa & Greenberg, 1989), which are generally abstract musical tones, auditory icons need not be novel sounds that must be learned by the users. They are based on natural sounds that a potential user would be familiar with from his or her daily life. For a brief comparison of earcons and auditory icons, see Houben (2002).

Auditory displays or sound graphs can be used to convey simple information through sound or to monitor multiple sets of data simultaneously. They have the benefit of being able to inform a listener of changing states in monitored signals without requiring visual attention. In the form of signaling systems, they can therefore be useful to blind people, but also just as useful to those with normal vision. Alarms are an example of a simple way of alerting a person, but displays that are more complicated exist such as an auditory workstation for an anesthesiologist from which the anesthesiologist can simultaneously monitor many types of information concerning a patient. For a review and more examples of auditory display technology, see Walker & Kramer (2004).

Machine recognition of sounds can be used for sound collection management (e.g., Brazil & Fernström, 2003), artificial intelligence in robotics, or for diagnostic purposes. Just as doctors may use sound to detect heartbeat irregularities and automotive mechanics use sound to hear engine anomalies, machines could potentially do the same. Okura (1999) described an example in which inspectors working in a Japanese canned-food factory tap canned goods with a metal rod in order to find defects. Such a task could conceivably be performed by a machine.

Knowledge of everyday sounds may also be used by sound designers who produce sounds for films, software, branding (e.g., radio or television jingles), and other applications (for a brief review, see Ballas, 2002). Appropriate design of sounds for branding purposes may allow the sound designers to portray the brand in a way that attracts consumers or suggests certain ideas to consumers (e.g., the strength of the brand's products).

An understanding of the perceptually important acoustic cues in everyday sounds can also be used to synthesize sounds for video games and other virtual reality environments (e.g., van den Doel & Pai, 1998; van den Doel, Kry & Pai, 2001). Although sounds can be prerecorded, using recordings has the inherent limitation that novel sounds may not be generated on demand. The information capable of being portrayed by sound is limited by the number of recorded sounds that have been stored. Synthetic sounds of sufficient fidelity may be superior to recordings because they can potentially cover a much wider range of situations than limited sets of recordings.

An obvious application of knowledge concerning the perception of everyday sounds is in assistive listening devices. The quality of life for people suffering from hearing impairments could presumably be greatly improved by hearing aids or cochlear implants made to be capable of conveying the important acoustic cues necessary for the perception of everyday sounds. Although this application is apparently an important one, the assistive listening device industry has primarily focused on improving speech perception instead of attempting to improve the perception of information provided by everyday sounds.

1.2 Thesis Organization

The aim of this thesis is to promote the idea of conducting ecologically valid research of the human auditory system and to generate basic knowledge concerning differences between normal-hearing and hearing-impaired listeners in the perception of everyday sounds. Investigations were performed that focused on producing knowledge concerning appropriate methods for investigating impact sounds and knowledge concerning the ways in which these sounds may be perceived by both normal-hearing and hearing-impaired people. An attempt was made to investigate ecologically relevant and perceptually relevant properties of a typical real-world sound event: an object impacting a surface. The design of the experiments emphasizes a need to maintain a holistic approach when studying hearing and hearing impairment. Some of the basic capabilities and uses of the auditory system, so quickly forgotten in traditional psychoacoustics, are stressed.

Chapter 2 provides background information concerning previous research that has been conducted on everyday sounds. The review is focused on impact sounds and on the information provided to listeners by impact sound events. The chapter also describes an information-based approach to perception, an “ecological acoustics”, which is commonly used when studying everyday sounds.

Chapter 3 provides some information on hearing impairment and hearing aids for those readers needing it. The chapter focuses on sensorineural hearing loss, the type of hearing loss from which the test subjects who participated in the experiments of Chapter 5 were suffering. Of interest to all readers may be an outline of hearing impairment as it relates to the perception of information, which is presented in Section 3.3. An information-based approach to correcting hearing impairment is proposed.

Chapter 4 presents the results of an experiment in which test subjects’ abilities to judge the lengths of wooden rods, solely based on hearing, was compared for three stimulus presentation methods. For the design of any experiment, a researcher must decide how to present the stimuli. This step may often not be given much thought, but it was hypothesized that the choice may have important ramifications of which the investigator should be aware. To test this, normal-hearing listeners listened to dowels of eight different lengths being dropped onto a linoleum floor. They did this while listening to 1) live presentation of stimuli, 2) headphone playback of binaural stimuli as recorded from an acoustic manikin, and 3) diotic presentation of monophonically recorded stimuli. A comparison of the length estimation accuracy for each of the presentation methods is presented. This chapter stresses the importance of presenting stimuli in an ecologically valid way when conducting research experiments. The results demonstrate that an experimenter's choice of presentation method can have an influence on the outcome of the experiment. Furthermore, a suggestion of a perceptually salient cue in the perception of length is made based on the outcome of the experiment.

Chapter 5 describes experiments with both normal-hearing and hearing-impaired subjects in which the subjects were asked to judge the lengths, materials, and heights from which a set of rods were dropped on the floor of a test room. It was hypothesized that hearing-impaired subjects would have more difficulty in judging these properties, which are highly relevant to people in their daily lives. Furthermore, it was desired to understand the influence of hearing aids in the perception of these properties. It was hypothesized that the hearing aids could contribute to improving perception in some ways, but may disrupt it in other ways. In order to investigate these ideas, twenty-four different stimuli were generated by dropping rods composed of two different lengths, two different diameters, and three different materials from two different drop heights. Performance for judgments of length, material, and drop height is compared for the two subject groups. Additionally, both the normal-hearing and hearing-impaired subjects performed the task with and without hearing aids so that the influence of hearing aids could also be examined. The normal-hearing subjects wore nonlinear behind-the-ear hearing aids specially fit to provide a small amount of gain, while the hearing-impaired subjects wore their personal hearing aids. A comparison of the results for each subject group, with and without hearing aids, is presented, as well as comparisons between unaided normal-hearing subjects and aided hearing-impaired subjects.

Chapter 6 briefly summarizes the findings that are further described in the body of the report and discusses applications of these findings. Suggestions for future work are also made.

Appendix A describes an experiment of more classical psychoacoustic style that was conducted as a part of this Ph.D. project. The intention of the experiment was to help bridge the gap between classical psychoacoustics and an ecological approach to auditory perception, but the results were not promising for bridging this gap.

Unfortunately, the results were not complimentary of classical psychoacoustics research. However, their inconclusive nature provides support for studying the human auditory system using an ecological approach. Because the description of this experiment would be somewhat disruptive of the flow of this thesis if it were placed as a part of the main text, it has been placed in Appendix A. A reproduction of the instructions given to test subjects who participated in the experiment is also included following the main report.

Appendix B contains a copy of the listening test instructions used for the experiment described in Chapter 4 along with example spectrograms and amplitude versus time plots of some of the stimuli used in the test.

Appendix C contains further details of the experiment described in Chapter 5, including the English and Danish-language versions of the instructions used in the listening test, technical information concerning the hearing aids used by the normal-hearing subjects, typical spectrograms and amplitude versus time plots of the stimuli used in the test, and a few additional results.

Chapter 2

Everyday Sounds

> *. . . sound perception is more than auditory sensation.*
>
> -Reinier Plomp (2002)

2.1 Information-Based Perception versus Sensation-Based Perception

Historically, most hearing research has been concerned with sensations and the belief that sensations are the “raw data” of perception. These can be referred to as sensation-based approaches to perception (Gibson, 1966). Loudness, pitch, timbre, and other sensory qualities have been exhaustingly (though not exhaustively) studied. Sensations arising from various, mathematically-simple, synthetic stimuli are well known, but little is known about how mathematically-complex real-world sounds are perceived. Attributes of sensation such as loudness vary constantly for real sounds, but how then does *constant* perception emerge from *constantly changing* sensations? It seems unfathomable, perhaps with good reason, that listeners could, without effort, make sense of so many constantly changing sensations. An information-based approach to perception suggests that perception is possible without sensory experience (Gibson, 1966).

Gibson (1966) argues that perception can occur without sensation. There is no reason to believe that the invariants that provide information to listeners are even open to analytical introspection, the process required to report sensations. He cites the vestibular system as an example of a perceptual system which constantly provides information but for which it is not clear that there are special modes of sensation (i.e., from the semicircular canals). Similarly, in sound localization experiments, listeners do not have two separate sensations resulting from the two ears. Simply a single experience of the direction of the sound occurs. This observation suggests that it may be a mistake to treat ears separately, and furthermore that sensation is not required for perception. If perception can arise without sensation as in these examples, Gibson argues that it may quite possibly also occur in other cases. This suggests that perception does not require sensation.

A major problem limiting classical psychoacoustics seems to be that a rigorous understanding of the auditory system in a purely physical way cannot account for its everyday functionality (Jenkins, 1985). The atomistic approach to understanding perception has not offered satisfactory solutions. A popular theory, for sensation-based perceptual approaches, is that a complex computational process is responsible for the reconstruction and interpretation of the noisy, inadequate components of the sensory input into a recognizable version of the original complex stimulus. The perceptual system must come up with a mental representation that is then compared to stored representations of sensory data (Noble, 1983; McAdams & Bigand, 1993; Carello et al., 2005). The best match is selected. It is suggested by some that such a process is controlled by an executive agency – a homunculus, or little man in the brain – that must perform the comparison of these representations.

Gibson (1966) has argued that the ear is not a tape recorder, just as the eye is not a camera, and that the formation of representations is unnecessary. Such representations would be akin to the image provided by a periscope – the homunculus looking at the reproduced image instead of the original. Instead, it is proposed that the active auditory system resonates directly to the information provided by events of the world (McAdams & Bigand, 1993). The goal of research following an information-based approach is not to determine how sensations are connected to representations, but how the sound of a closing door is discriminated from other sounds. Some of the burden previously placed on the nervous system has now been placed on the signal itself. A catalog of sensations as a preliminary requirement for perception is no longer required (Gibson, 1966). The differences between these theories of higher order processes are a hotly debated topic (Neuhoff, 2004c). They will not be discussed further, as they are not the focus of this thesis, but this should serve to point out that alternative theories exist, which do not run into the same roadblocks encountered by sensation-based theories of perception.

Traditionally, psychoacousticians have studied the transduction of sound that is performed by the ear and some of the higher level experiences elicited by proximal stimuli (Gaver, 1988). Sound, as vibrations around the head (or only a single ear), have been considered as the input to the auditory system in most cases. Consideration for the source of the sounds has not been made. The results of such studies may be considered valid in the sense that they properly describe a response to an acoustic input. It is an attractive feature of this approach that descriptions resulting from such studies are (presumably) applicable no matter what source has created the acoustic input. However, if the acoustic inputs studied are seldom observed outside of laboratories, or if they are devoid of information, their usefulness may be limited. A useful study of auditory perception includes consideration for the information that the system is designed to gather.

Gaver (1993a) suggested that all sounds can be heard in two different ways. One could listen to the sensory qualities of the sound, the loudness, pitch, timbre, and so on, but one could also listen to the sound in terms of its source – the event and objects involved in the event – that produced the sound. He called the first kind of listening *musical listening* and the second kind *everyday listening*. Everyday listening is the type of listening that people do in their daily lives. Musical listening is the type of listening done in most psychoacoustic laboratories, when tuning a guitar, or when trying to describe a sound that cannot be recognized. Musical listening has also been referred to as *reduced listening* in order to avoid the potential misinterpretation that the expression concerns only music (e.g., Schaeffer, 1966; Casey, 1998). In experiments described later in this chapter (e.g., VanDerveer, 1979), test subjects have been found to listen to sources, and not to sensory qualities, when simply asked to describe what they hear. Research involving sounds that do not carry information (e.g., pure tones) forces test subjects to listen musically. However, at least for real-world sounds, it is usually possible to listen in either way. Listeners can introspect in an attempt to analyze the sensations produced by a sound entering their ears, but they can also, perhaps more easily, hear the entire "complex" sound as a whole. In the end, it is the information that is important. The sensations are incidental (Gibson, 1966).

Sound carries information about the world. Information about events in the world, along with information about the environments in which they occur, is provided to those within audible range. Sound provides news of an event (Jenkins, 1985). An information-based approach to auditory perception considers the ears as active parts of a system used to gather information in its environment. In contrast to a sensation-based approach, the ear as a passive receptor is not the focus. Instead, the focus is on the ways in which the entire auditory system, including the muscles capable of orienting and moving the body to the source of sounds, acquires and uses information. The focus is on *everyday listening*, not musical listening. An information-based approach suggests that listeners attempt to detect the invariants present in sonic events, despite constant changes in sensations (Gibson, 1966).

Jenkins (1985) described an example with speech: without the effort required to describe a single sensation, listeners can simultaneously gather a great deal of information based on a speaking voice. One can typically hear the language, dialect, and/or accent of the speech, the emotion of the person producing it, the direction and distance from which the speech signal originates, the age, gender, and perhaps the identify of the speaker, whether the words are being whispered, shouted, or sung, all of this along with the possibility of also simultaneously understanding the meaning of the words. Only in extreme cases, or when requested to do so, do listeners report on the pitch or loudness of a voice.

The goal of an information-based approach to auditory perception is to understand which features of sound carry information to listeners about events in the world (Gaver, 1988). While a sound may be physically analyzed in a great number of ways, a focus on the properties of the sound important for the perception of information is necessary. Cognitive research concerned with the experience provided by the auditory system has commonly occurred separately from psychoacoustic research (Plomp, 2002). An ecological version of psychoacoustics, combining classical psychoacoustic and cognitive approaches, is in order (e.g., Neuhoff, 2004b). An approach suggested by Werner & Liebold (2004) is that an ecological psychoacoustics should first identify the acoustic information provided by sound events and determine what part of this information is used by listeners. This knowledge can provide the foundation for more in-depth research.

The classical emphasis on how the ear works instead of what the ear hears should be reconsidered (e.g., Gibson, 1966; Gaver, 1993c; Rosenblum, 2004), but neither the approach of only studying the physiology of the ear nor the approach of only studying what the ear hears will likely be successful at providing a global understanding of the auditory system. If man is to one day be better equipped to address the problem of hearing impairment, to produce more useful auditory feedback for man-machine interfaces, to improve computational speech and sound recognition algorithms, and to simply understand the role of sound in the lives of humans, a holistic approach which considers *why*, *what*, and *how* is required (e.g., Gaver, 1993a,c). This thesis is primarily concerned with the neglected issue of *what* the ears hear. It is concerned with a study of the useful sensitivity of the auditory system, not with the sensitivity of its receptors.

2.2 Information From Everyday Sounds

Gibson (1966) defined a framework for studying perception in an ecologically relevant manner. He addressed the uses of many perceptual systems, including the

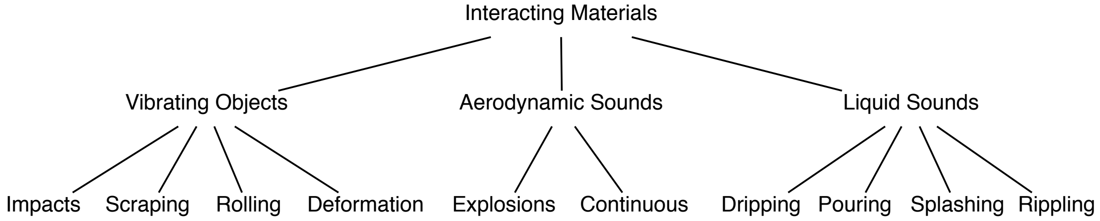


Figure 2.1: Hierarchical categorization of everyday sound producing events as defined by Gaver (1993c).

auditory system, but his analytical and experimental focus was on visual perception (e.g., Gibson, 1979). Ecologically motivated investigations of everyday sounds (e.g., VanDerveer, 1979; Gaver, 1988), of which the first significant research occurred less than 30 years ago, initially involved *protocol studies* in which listeners heard a variety of sounds and were then asked to identify what they heard. Similar studies continue today (e.g., Gygi, 2001; Shafiro, 2004b), and are still producing results of fundamental importance to describing everyday sounds. By their nature, these types of studies generally do not provide much information about perception of specific types of sound events, but rather they describe listening skills that may be common to sound perception in general. Along the way, a great deal of insight is typically provided concerning which types of sound events may be worth exploring more deeply.

In part of his work in everyday sound research, Gaver (1993c) defined a basic categorization system for everyday sound events. Three broad classes of sound events were defined which could all be roughly described as material interactions. Within each of these classes, there are more specific types of simple sonic events. His original diagram has been reconstructed, slightly modified, and is shown in Figure 2.1. The categorization system is founded in both physics and perception. In sound identification studies in which a large variety of sounds are presented to test subjects who are asked to describe what they hear, there are generally no confusions between the three broad classes of the sound events at the middle level of the diagram. Confusions that occur are generally between the subclasses of the three main categories (e.g., VanDerveer, 1979; Gaver, 1988).

The basic level sources defined in each category are descriptions of simple events. Temporal patterning of the events may modify them to produce more complicated events, and similarly, compound and hybrid events may exist which can best be described as a combination of basic level events. Each of the basic level events, and thereby more complicated events, is capable of conveying information, and many of these events have been shown to convey this information to human listeners.

2.2.1 Aerodynamic and Liquid Sounds

Of the three main categories of the classification system, aerodynamic and liquid sounds have received the least attention. According to the classification system, aerodynamic sounds can be grouped into two categories: those created by sudden pressure changes (explosions) and those created by continuous aerodynamic events such as wind. Research concerning the perception of these kinds of sounds is scarce.

As with aerodynamic sounds, liquid sounds may be separated into categories divided by how stationary the signals are: dripping, in which the sounds are separate impulsive events, and pouring, rippling, and splashing in which there is greater continuity tying the events together. Few experiments have investigated liquid sounds, but one found that test subjects were able to utilize acoustic information provided from liquids in containers (Cabe & Pittenger, 2000). In an experiment using monophonically recorded stimuli presented from a loudspeaker, test subjects were found to be able to categorize whether the water level in a container was constant, rising, or falling, based on sound. In a second experiment, blindfolded subjects were asked to fill a container to either drinking level or the brim of the container. The test subjects were clearly able to distinguish between the two target levels. Container size and water flow-rate were varied in an additional experiment in which both blindfolded subjects having normal vision and subjects who were blind were asked to fill containers to their brims. Most subjects were successful at stopping the filling near the brim of the container, but it was found that the subjects did not all make optimal use of the acoustic information available. No difference was found between the performance of blind and sighted participants.

A theoretical acoustic analysis indicated that the fundamental resonant frequency of the event was sufficient to indicate whether a container was being filled or emptied. A high fundamental resonant frequency was found to indicate that the container was near full, and the rate at which the frequency increased was found to indicate how far the container was from being full. A rapidly changing fundamental frequency indicated that the container was very close to being full.

In a final experiment, blindfolded test subjects listened to live stimuli of containers being filled at different rates. The filling of the containers, which began at different initial fill heights, was abruptly stopped during each fill, and the subjects were asked to indicate when the container would have been full if the filling had continued. Subjects were reasonably good at predicting the point at which the containers would have been filled; however, the authors were cautious in claiming this result was meaningful. In summary, the experiments described by Cabe & Pittenger (2000) demonstrated that listeners are able to make use of information provided by liquids filling containers in order to control the height of the liquid. Additionally, the research indicates that sound may also provide information concerning the approach of an event, or “time-to-arrival”. This kind of information will briefly be discussed beginning on page 54.

2.2.2 Vibrating Objects

Of the three main categories shown in Figure 2.1, the category of *vibrating objects* has received the most attention from everyday-sound researchers. In general, these types of sounds may convey information concerning size and material. Sounds of each subtype may convey further information. Because they are the focus of this project, impact sounds will be described in the most detail. However, other notable experiments, and those that are relevant to this thesis, will also be mentioned.

Rubbing and Scraping Sounds

Scraping may convey information about the materials and textures of the objects involved in the event, the force with which the two objects are scraping against each other, and information about the velocity and acceleration of the two objects (Gaver, 1993c).

Halpern, Blake & Hillenbrand (1986) investigated scraping sounds, like that of fingernails beings scraped across a chalkboard, in order to determine the acoustic qualities that make the sounds unpleasant. Monophonic recordings of various sound events were made, some of which were scraping sounds and some of which were not. The authors asked subjects to rate the pleasantness of the sounds. A slate surface being scraped with a three-pronged garden tool was confidently found to be the most unpleasant, closely followed by other scraping events: two pieces of Styrofoam scraping against one another, scraping metal, and scraping wood.

Further experiments were then done to try to identify whether the spectral content or temporal structure were responsible for unpleasantness. A “particularly aversive” version of the garden tool being scraped across slate was low-pass and high-pass filtered with various cutoff frequencies. Unexpectedly, it was found that amplitude-normalized high-pass signals, with cutoff frequencies of 4 kHz and higher, were significantly less bothersome than any other versions of the signal.

Next, a version of the signal was constructed of which the temporal fluctuations had been removed. Similar synthetic versions of the signal with and without the temporal fluctuations were created using sinusoids mimicking the first three prominent harmonics of the real sounds. The original recordings, in both the original form and the form of which the temporal fluctuations had been removed, were judged more unpleasant than the synthesized signals. The removal of the temporal fluctuations from the recording was not found to influence the unpleasantness of the signal. The results suggested that the frequency content of the stimulus was responsible for the unpleasantness of the sound event. While a specific characteristic of the acoustic wave could not be specified to cause the unpleasant character of the sound, the high frequencies were found not to contribute to it.

Pleasantness is not a measure of information, and for that reason, the above experiment could be said to be not asking an ecologically relevant question. The authors pointed out however that the sound produced by the scraping of the three-pronged garden tool across a slate surface is very similar to warning cries produced by macaque monkeys. If the sound is similar to a biologically based warning signal, then it could be providing information to the listener. Nevertheless, the stimuli used in this experiment were ecologically founded.

Rolling Sounds

Rolling may convey information about the materials and textures of the objects involved in the event. In addition to details concerning the regularity of the rolling, information about the circumference and weight of the rolling object may also be available. Angular speed or linear speed may be deducible from this information (Gaver, 1993c).

Fowler (1990) conducted experiments in which listeners heard monophonic recordings of a 2.54 cm (1 in) diameter steel ball rolling on steel ramps composed of two straight segments. Listeners were asked to judge the steepness of the first segment – a ramp. The ball was released from the top of the first segment. It rolled down the first segment, which was covered in sandpaper to make it quieter, and onto the second segment composed of steel. The steepness of each segment was variable. The first segment of the track was set to between $10\,^\circ$ and $50\,^\circ$, relative to horizontal, while the second portion of the track was either flat or inclined at an angle of $23.5\,^\circ$. For the track in which the second portion was flat, the ball was caught after falling off the end of the track.

Information concerning the steepness of the initial ramp was theoretically available in the duration of time the ball spent on the second segment, but this information was not alone sufficient for correct judgments. For tracks in which the second segment was flat, a steeper initial ramp resulted in a shorter time spent on the second segment of the track. Shallow initial ramps led to the balls spending longer periods of time on the flat, second portion of the track. For tracks in which the second segment was inclined, the *opposite* was true. A steeper initial segment resulted in a longer time spent on the second segment of the track, because the ball would go up the incline and then come back down. A shallow initial segment resulted in a short amount of time being spent on the second part of the track.

Monophonic recordings from the first and second segments of the track were split and spliced together to form synthetic events. Test subjects were asked to judge only the steepness of the first part of the track. Results indicated that the subjects used information from the sound of the ball on the second part of the track to make judgments of the initial ramp steepness, and that they could do so regardless of whether a short time spent on the second segment indicated a steep initial ramp or if the opposite was true. It was later emphasized by Carello et al. (2005) that had only flat tracks been used in this experiment, the authors may have been led to incorrectly hypothesize that the duration of time the ball spent on the second segment of the track was solely responsible for judgments of steepness of the initial ramp. This would of course, been in error, as was neatly demonstrated that listeners must be provided information concerning steepness in other ways.

An additional experiment was conducted in which the recording of the ball rolling from the 10° sloping ramp onto the flat second segment of the track was modified. It was modified so that the time spent on the second segment of the track was shortened to be equal to the time that was spent on the second segment of the track when the ball initially rolled down the 50° ramp. The results from this experiment indicated that listeners were still able to use information from the sound of the ball on the second part of the track to make judgments of the initial ramp steepness. Further evidence was therefore provided, which showed that duration was not the only cue used to make judgments of the ramp's steepness. As an example, the authors pointed out that the revolution speed of the ball could also possibly be used by subjects to perform the task. Of general significance, these experiments indicate that information from an acoustic event is not only informative about what is occurring at the time it is received, but it can also be informative about previous developments of the event.

Houben (2002) investigated the ability of listeners to hear the sizes and speeds of rolling balls (see also Houben, Kohlrausch & Hermes, 2001, 2004). In the experiments, monophonically recorded stimuli and manipulated versions of the stimuli were presented to test subjects who were asked to judge the sizes and linear speeds of wooden balls rolling on wooden plates. In this case, the sound primarily radiates from the wooden plate (sounding object), but is excited by the ball (non-sounding object). Stimuli were presented in two-interval, two-alternative forced choice tasks. It was found that subjects could generally categorize size and speed differences even when both were varied simultaneously. There was, however, an influence of each physical property on judgments of the other. It was noticed that when the stimuli were normalized to have the same sound pressure level, judgments of speed were more likely to be incorrect. Furthermore, it was found that spectral cues dominated the judgments of size and speed, but this could have been simply because the stimuli were selected to contain a minimum of temporal cues (e.g., the tracks were made to be as smooth as possible). In an artificial case, temporal amplitude modulation was added to the stimuli. Speed judgments were strongly affected. This suggested that one should use caution before concluding that judgments were influenced by only spectral or only temporal cues. Apparently, multiple cues potentially contribute to the perception of size and speed.

The experiments of Houben suggest that humans are capable of hearing the size and speed of rolling balls. However, it is questionable whether these experiments have demonstrated the way in which listeners do this in the real world. It is likely that the cues found to influence the perception of size and speed in these experiments do actually influence the perception of size and speed in the real world, but the degree of their influence is not clear. Likewise, there may be other cues, absent from the experiments, which are just as important. The fact that this is possible was demonstrated by the author himself, as it was found that speed judgments were worse when sound pressure level differences between the recordings were removed. Additionally, spatial cues were missing from all of the stimuli due to the use of monophonic recordings. Spatial cues could in many cases have an influence on the perception of size and speed. The most easily explained example of this is that if one is able to perceive a rolling ball's change in position over time, then the listener also has knowledge of its speed. In conclusion, the removal of spatial cues and the removal of temporal variation, both of which may exist in natural settings, have perhaps hidden acoustic cues that may be used for size and speed perception in a listener's everyday life.

Impact Sounds

Most of the research of everyday sounds has focused on impact sounds. In general, impact sounds may convey information about the force of the impact as well as the vibrating object's material, size, surface hardness, and configuration (Gaver, 1993c). Impact sounds are everywhere. Examples include a door closing, a drinking glass being set on a table, a kicked ball, a closing book cover, a visitor knocking at the door, a housefly flying repeatedly into a window, footsteps, typing, stapling, and many more. Hundreds if not thousands of impact events occur within audible range of a listener on a daily basis. A few experiments, some of which are presented below, have been conducted to determine if people are in fact capable of obtaining information from the multitude of different kinds of impact events. This Ph.D. project has focused on investigating the types of information and the accuracy of the information that listeners are capable of obtaining with their sense of hearing. Because of this focus, emphasis has been placed in describing experiments that have studied *what* people are capable of hearing as opposed to *how* they hear it, but experiments concerning how people hear are also touched on.

1980's Warren, Jr. & Verbrugge (1984) found that listeners could discern, with very high accuracy, bottles that broke when dropped compared to bottles that bounced when dropped. In an attempt to understand the cues important for listeners to make the distinction between breaking and bouncing, the authors synthesized breaking and bouncing events using temporally manipulated recordings of bounces. The authors encouraged a test methodology of "analysis and synthesis", where physical and acoustic analyses were used to help characterize events and produce synthesis algorithms, which in turn are verified through listening tests and modified as necessary.

Separate monophonic recordings were made of four broken pieces of glass being dropped and then bouncing. To simulate the bouncing event, the four recordings were manipulated so that the onsets of the bounces of each of the four separate recordings each occurred in unison. The four separate recordings were then combined to form one event in which each piece bounced at the same times as the other pieces. The breaking event was synthesized similarly, using the same four original recordings, but with the bounces adjusted to occur asynchronously.

Listeners were able to discriminate between the resulting "synthetic" breaking and bouncing events nearly as well as the original monophonic recordings of breaking and bouncing events. The authors therefore concluded that the temporal envelope of the signal, the primary aspect that had been manipulated in the creation of the synthetic events, was important to listeners in discerning these two events. The authors noted that although the synthetic breaking and bouncing events were clearly discriminable, the synthetic events sounded more as if they involved metal objects than glass. They attributed this to the spectral properties produced by the sounds of the four particular separate pieces of glass that were combined to generate the synthetic breaking and bouncing events. The synthetic stimuli had not preserved the original spectra. Their findings suggested, as VanDerveer (1979) originally proposed, that spectral information may convey information about material, and that the temporal envelope structures information about the event.

Warren, Jr., Kim & Husney (1987) tested the abilities of people to bounce rubber balls to a target height in a multimodal experiment. Prior to the subjects bouncing the balls, they were given one of four possible previews of the balls: 1) no preview, 2) an auditory preview of the ball bouncing, 3) both an auditory and visual preview of the ball bouncing, and 4) a chance to dribble the ball twice before the test bounce. In all cases, unavoidable haptic[1] cues may have been available to the test subjects simply through touching the ball before the test bounce. The balls used in the experiment had different elasticities, but equal weights. Subjects performed best for the condition in which they were able to dribble the ball themselves before the trial, but subjects performed significantly better than the no preview condition for all other conditions. No statistically significant improvement was found between the auditory and auditory/visual preview conditions. In another experiment presented in the article, simulations of a bouncing ball were generated for three conditions: visual-only, auditory-only, and auditory/visual combined conditions. Bounces were conveyed acoustically by a short, low frequency tone. It was found that the information available to subjects concerning period between bounces was equivalent for the auditory and visual modalities. The results from this experiment again demonstrated that temporal patterning is informative about events.

Repp (1987) examined hand clapping. He found that the identity of subjects whose handclaps were recorded and played back for other subjects were not reliably able to be discerned, nor were the genders of the clappers. Only self-identification was found to be good. On the other hand, listeners could discern between hand configurations. Spectral differences between configurations were believed to be responsible for this ability. The author concluded that the results provided evidence that listeners are capable of hearing changes in the states of the objects involved in sound events.

Wildes & Richards (1988) made a physical model in order to demonstrate that specific loss, a material property related to measures of the internal friction of the material, can be used to estimate the width of a resonant peak of an impact sound and the rate at which this resonance decays. Although no listening tests were conducted, the authors demonstrated that information describing material could be present in an acoustic waveform. They suggested that it was feasible to determine the theoretical possibilities and limits of material perception. It was proposed that their work could be used to guide auditory experiments, helping to confine the experiment to those tasks that are physically possible. However, their model only examined two potential acoustic cues. While these cues may in fact be the only cues responsible for conveying material, a cautious researcher may wish to

---

[1] "Of or relating to the sense of touch" (Pearsall, 1999).

verify this assumption. It may be dangerous to allow their model to be a complete guide of what one tests and what one does not test. Their work may, however, be helpful in explaining results.

It was later pointed out by Tucker (2003) that an assumption made by Wildes & Richards (1988), that the relationship between decay time and frequency are inversely related, may not be very accurate. Other experiments (Krotkov, Klatzky & Zumel, 1996; Femmam, M'Sirdi & Ouahabi, 2001) showed that the relationship between these two parameters was better described by a polynomial. These discrepancies draw attention to the fact that the assumptions made when producing models may be critical in the accuracy of the model. If a research goal is to understand how well listeners can hear the nature of a material or to understand how listeners hear the nature of a material, it may therefore be risky to trust mathematical models used in the analysis or synthesis of the signals. Assumptions may result in details of important perceptual consequences being ignored. It seems that an initial test of a model should be for the model to be able to recognize materials from binaural recordings at a rate at least better than the best human performance found.

Gaver (1988), as a part of his Ph.D. thesis, performed listening tests in which listeners were asked to identify the material and length of bars struck with a mallet. Recordings were presented to subjects in which the materials and lengths of the bars were varied simultaneously. Discrimination between metal and wood was nearly perfect. Relative length was found to be perceivable independently of the material of the bar, but there was a tendency for subjects to rate the shortest and longest metal bars at more extreme lengths (shorter and longer, respectively) compared to the wooden bars. There were also somewhat large individual differences in length estimations, with some subjects performing very well and others very poorly. While the relative lengths of the bars was fairly accurately represented for the group average data, the individual subject data suggested that subjects were fairly inaccurate at judging the lengths of the bars. Gaver (1988) therefore concluded that subjects were better at judging the materials of objects than the lengths of objects.

1990's Lutfi & Oh (1997) conducted listening experiments to determine the capability of normal-hearing listeners to hear the material of synthesized struck-clamped bars. The synthesized "materials" they selected for comparison were chosen because of a theoretical ability to discriminate between their frequency components, amplitudes, and decay rates (Wier, Jesteadt & Green, 1977; Jesteadt, Wier & Green, 1977; Van Heuven & Van Den Broecke, 1979). They compared, for example, four metals to one another: iron, silver, steel, and copper. They used synthesized signals for their test, stimuli that were created by theoretical acoustical models. These synthetic stimuli were truncated to 400 ms in length, even though the uncut versions had audible information for "several seconds after the bar was struck." The authors only synthesized the first three harmonics of each hypothetical bar, as those were believed to be the only partials that were typically audible.

Using eight extensively trained musicians as test subjects, they found that listeners did not take advantage of the cues available to them in order to do this task. The resulting performance was "far less than ideal." Lutfi & Oh (1997) expressed in their findings that they believed the views of Gibson (e.g., 1966, 1979) were optimistic and that their results were in better agreement with the "less optimistic" results of Wildes & Richards (1988). The authors pessimistically concluded that listeners were poor at discriminating materials. If considered as ecologically valid, their findings suggest that there are limits to human abilities to discern materials. The results indicate that a normal auditory system may not be capable of differentiating between different types of metals or in other materials that are similar in sound. Listeners did not take advantage of all acoustic information made available to them, instead relying primarily on frequency. They used amplitude and decay rate only secondarily. As Gaver (e.g., 1988) had previously shown though, listeners are perfectly capable of making categorizations of materials which are grossly different.

An alternate explanation to that provided by Lutfi & Oh (1997) could simply be that the authors did not test material perception – they used no real materials. They tested the way in which listeners try to map the characteristics of acoustically impoverished signals to things they know from everyday life. There is no reason to expect listeners would be good at this task – at listening to truncated signals from mathematical models that are lacking in complete information. Other work from the same lab (Lutfi & Oh, 1994) attempted to validate their models by comparing spectrograms of synthesized versions of a tuning fork to real versions of a tuning fork[2]. They found that only with care could they strike the tuning fork in a way that produced a spectrogram that was nearly identical to a spectrogram of the synthetic signal. The connotation in the article was that the variation present between strikes of the tuning fork was a problem, but they did not provide an explanation as to why this should be the case. It was later pointed out by Carello et al. (2005) that Lutfi & Oh (1997) seem to have believed that simultaneously changing multiple variables during the test and testing inexperienced listeners to judge parameters would make their tests more difficult. In experiments by Kunkler-

---

[2] The sound of a struck tuning fork is certainly one of the most acoustically- and informationally-impoverished sound events possible.

Peck & Turvey (2000) and others, these ideas were later shown to not necessarily be true. It has been suggested that the added natural variation makes it in fact easier for listeners to extract the relevant invariants from the stimuli. The choice of stimuli in the experiment of Lutfi & Oh (1997) appears to have been guided by classical psychoacoustics and not by ecological value.

The first author later explained that error-free performance was not even possible in the experiment (Lutfi, 2001). He stated in retrospect that the study required "a decision between two bars of the same material in different relative concentrations." This fact was believed to play a role in the outcome of the experiment, but he wrote that the results were not necessarily "specifically related" to this problem.

Lakatos, McAdams & Caussé (1997) tested the abilities of subjects to discriminate between the cross-sectional shapes of metal and wooden bars. In pilot studies, metal and wood, as materials, were found to be almost perfectly discriminable. The stimuli were presented to test subjects from AB stereo recordings of free-hanging bars being struck by mallets at their centers (center of bar in relation to the length and width). All bars were 30 cm in length. Subjects responded by selecting one of two alternative cross-sections using a computer interface. For metal bars, subjects were accurate in their choices 77.6% of the time, while the results for wood were slightly worse with 73.8% of responses being correct. The authors attributed the worse performance with wood to the shorter decay time of wood (samples truncated to 250 ms vs. 2500 ms for metal). Best performance among all bars was found to be correlated to large differences between the width and height ratios of the two alternative bars.

An acoustic analysis of the signals was also performed. Although the theoretical frequency components had to be searched for at times (and in some cases were completely missing), the frequency components in general were highly correlated to both the width/height ratios and to the judgments made by subjects. The authors summarize that listeners have at least a basic ability to distinguish the cross-sectional dimensions of impacted objects, but that they also used only a part of the acoustic cues available to them in their decisions.

Houix, McAdams & Caussé (1999) struck metal bars of varying cross-sectional dimensions in five different positions in order to excite various vibratory modes (thickness-related transverse modes, width-related transverse modes, and torsional modes). The authors asked listeners to categorize a complete set of recordings of five strikes on six bars. In separate experiments, subjects were asked to group the sounds that "sounded similar" and those that could have come from striking the same bar. The listeners categorized the 30 recordings, but the results indicated that they were not able to clearly identify the bars. The authors concluded from the categorizations that pitch was the primary cue used by the subjects to perform the sorting task. The subjects were found to not take advantage of all of the acoustic information available to them.

Freed (1990) asked listeners to listen to recordings and judge the hardness of mallets used to strike cooking pans. The task was therefore one in which listeners judged a property of the non-sounding object (the mallet), not the object which resonated most of the sound (the cooking pan). Subjects were successfully able to scale the hardnesses of the mallets, irrespective of the size of the pans that were struck. Differences between mallets are well known to be of importance to musicians (e.g., Fletcher & Rossing, 1998), but whether mallet hardness itself is directly perceivable had not previously been demonstrated. The author identified four acoustical parameters of the attack portion of each strike and found that they were perceptually correlated to the results: a mean measure of the overall sound level, the slope of this level with respect to time, a measure of the mean spectral centroid[3], and a measure of how this centroid change with respect to time. The acoustic analysis was later criticized by Lakatos et al. (1997), as the acoustical descriptors were not invariant across the pans even though the mallet hardness judgments were. The results were most interesting though because subjects were able to judge a property of the mallet separately from that of the sounding object itself, as in this case, the pan is the primary resonator of sound.

Li, Logan & Pastore (1991) tested listeners' abilities to categorize the gender of human walkers based on recorded walking sounds. The task was therefore again one in which listeners judged a property of the non-sounding object – the walker – who himself or herself did not radiate sound. Males and females were recorded walking on a hard surface, and subjects were asked to categorize presentations of four steps as being produced by a male or a female walker. Subjects generally performed quite well. Spectral properties were found to play an important role in categorizing male and female walkers, but temporal factors were not found to provide information concerning gender. Fast walkers were more likely to be judged as female, but the actual walking paces did not significantly differ between the male and female walkers. It was proposed that the stereotype that high pitch is associated with femininity was derived from generally true differences. The authors suggested that the size of a source may be indicated by its fundamental frequency. The types of shoes worn by the walkers also influenced gender judgments. Judgments were worsened when female walkers wore male shoes. Among various anthropomorphic measurements of the walkers, many of which were highly correlated with gender, the height of the recorded walker was the best correlated to the judgments.

---

[3] A measurement of the distribution of spectral energy.

Heine, Guski & Pittenger (1993c) tested the ability of listeners to hear the number of balls being simultaneously dropped on a hard surface (for a shorter report of the same experiment, see also Heine, Guski & Pittenger, 1993a). The authors performed an acoustic analysis to demonstrate that there are differences in the temporal structures and peak sound levels of the acoustic signals. The A-weighted peak sound level increased as a quadratic function with an increase in the number of balls. Temporally, the total duration of the event was greater for a larger number of balls, and the time between peaks in the signal were shortened. In principle, there were therefore at least three different forms of acoustic information that could be available for discrimination. Test subjects performed the test by listening to binaural recordings, which had been recorded with an acoustic manikin, of between one and ten steel balls being dropped onto a wooden surface. Subjects underestimated the total number of balls in most cases, but the resulting estimated number of balls were fairly well correlated to the actual number of balls ($r = 0.73$, $p \leq 0.001$). Although they were unable to predict the precise number of balls (except when there was only one ball), subjects showed an excellent ability to detect an increase in the number of balls. The authors suggested that the auditory system may not be optimized for counting objects but for performing gross categorizations in such tasks. Other perceptual systems may be better for counting the precise number of objects involved in an event. Other researchers have made similar conjectures concerning the precision and uses of the auditory system (e.g., Popper & Fay, 1997; Grassi, 2005).

Pittenger, Jordan, Belden, Goodspeed & Brown (1997) examined whether or not information for the perception of size was available from shaking and stirring events, in an experiment testing both auditory and haptic perceptual systems. Test subjects or the experimenter either stirred or shook balls in containers, and the subjects were asked to rank the containers based on the size of the balls in the containers. The balls ranged in size from 3.6 to 10 mm in diameter, the smaller of which were steel, and the larger of which were plastic. In the haptic condition, the subject stirred or shook the containers while masking noise was used to deafen the auditory sense. In the auditory condition, subjects listened to the experimenter stirring or shaking the containers. Finally, in a combined condition, subjects themselves stirred and shook the containers, without any masking noise. The results for the haptic condition showed that subjects were able to perform the task at a level better than chance, and for the auditory condition, the subjects were nearly perfect at ranking the sets of seven (plastic) and eight (steel shot) containers. The worst performance occurred for stirring the smallest steel balls, which were different in diameter by only 0.25 to 0.5 mm. The combined auditory and haptic condition showed no benefit of having haptic information in addition to auditory information.

Pittenger & Mincy (1999) performed an investigation similar to Pittenger et al. (1997), this time in which subjects were asked to rank the sizes of granules inside opaque containers. The granules ranged in width from approximately 64 to 665 μm. As in the experiments of Pittenger et al. (1997), the subjects' abilities to sort the containers by the size of the granules inside, was tested for auditory, haptic, and combined conditions for both shaking and stirring. As with the larger balls of the previous experiment, subjects were generally quite good at this task when performing it based on acoustic information. Performance in the auditory condition was better for shaking. Haptic performance was very good in the stirring condition, but poor in the shaking condition. For stirring, but not for shaking, combined auditory and haptic information improved results beyond either modality alone. In summary, yet another experiment has shown that size is perceivable from acoustic cues, but the authors stressed that multiple measures of how well perceptual systems obtain information from their environments should be made. Accuracy measures alone should not form the only investigation.

Although it may be of small importance, it is unclear if the stimuli used in the above task should strictly be considered an impact event, and therefore whether or not a description of this work should fit in this category of the classification system being used here (the classification system of Gaver, 1993c). The stirring and shaking of containers containing small granules creates thousands of tiny impacts, but may also involve scraping for example. The event may alternatively be considered as a compound or hybrid event. For this literature review however, the experiment description has been placed in the impact sound category.

Anderson, Peck & Carello (1997) presented a preliminary report concerning experiments conducted in which listeners were asked to estimate the lengths of wooden dowels dropped onto a surface. A regression of the logarithm of perceived rod length onto a function involving the principal moment of inertia of the rods ($r^2 = 0.97$) showed that length perceived by sound is constrained by the same physical quantity (moment of inertia) that had previously been found to constrain length perceived by wielding rods by hand (Solomon & Turvey, 1988; Fitzpatrick, Carello & Turvey, 1994). The authors cited these results as evidence of the nervous system using physical invariants discovered by interaction in the physical world.

Carello, Anderson & Kunkler-Peck (1998) provided further details of the previously reported experiments on auditory perception of the length of dropped wooden dowels (Anderson et al., 1997). In two experiments, subjects listened to pine dowels of equal diameter being dropped onto surfaces (a linoleum floor in one case, plywood in another). Stimuli were presented live. The rods fell on the right-hand sides of the subjects, who were separated from the stimuli by a Styrofoam screen. In each of the experiments involving two different diameters between the experiments, seven rod lengths (30 to 120 cm, 1.27 cm diameter in Experiment 1; 10 to 40 cm, 0.32 cm diameter in Experiment 2) were presented to the subjects who responded by positioning a moveable surface at the desired distance from the edge of a desk. The subjects were asked to position the moveable surface in such a way that it could just be reached if the rod were extended from the edge of the desk to the surface.

The results from the combined data indicated that perceived length was tightly coupled to actual length. While estimates were not perfect, the mean data showed that the average data of the subject was a nearly perfect rank of the rod lengths. The length estimates were also impressive in their absolute accuracy. Individual subjects also showed similar results, albeit some better and some worse. The results were comparable to those previously found for wielding non-visible rods.

Although it was unclear from the short article exactly how it was conducted, a basic acoustic analysis was performed to determine whether signal duration, amplitude, or spectral centroid could account for perceptual performance, but none of the independent regressions showed promising results. As previously explained for the experiments of Anderson et al. (1997), the inertia tensor, a measure of the rod's resistance against rotation, was suggested as having a strong relationship to perceived length. However, actual length was found by regression analyses to have just as strong of a relationship. The authors proposed that the expression involving moment of inertia was more logically founded as a better representation of the rod as a mechanical structure. Later experiments would show that Young's modulus of elasticity in addition to the inertia tensor could additionally account for rods of different densities (Carello et al., 2005). Rosenblum (2004) suggested that these quantities were certainly specified by higher-order descriptions of the acoustic dimensions that, on their own, may have previously failed to account for length perception.

2000's Cooper, Janovicz & Carello (2001) continued with previous investigations concerning the perception of rod length (e.g., Carello et al., 1998), with a new experiment in which pine rods were struck with wooden mallets. Five rod lengths, from 30 to 90 cm, all with 1.2 cm diameter, were struck at their centers of mass in two conditions: when freely suspended and when clamped to a hard surface. Blindfolded subjects reported perceived lengths by positioning their hands in such a way that they could just hold the rod, end to end. The experimenter measured the distance between their hands. Twelve of the subjects performed the test without feedback, but seven additional subjects performed it with visual feedback of the actual rod length following each trial. Rod length, feedback, and an interaction of rod length and the way in which the rod was supported were found by a mixed-design analysis of variance to produce significant effects.

As had been previously found by Carello et al. (1998), subjects were generally able to perceive the relative sizes of the rods, but in this case, the average perceived lengths were compressed relative to the actual lengths. The estimated lengths of the shortest rods were fairly accurate over all, but the estimates of the longer rods' lengths were significantly lower than the actual lengths. Although it was not suggested in the short article, it is possible that this could be due to a bias resulting from the response method. Feedback was found to significantly affect the results. Those receiving visual feedback had responses that were more accurate on average, but the variation between their responses was not improved. In other words, their fixed error was lower, but their variable error was not. The authors had no explanation for the fact that there were differences between length predictions for some of the rods that depended on the way in which the rods were supported, but the sound produced in such cases is obviously different.

Wagman (2003a) performed an investigation to determine if training or feedback had an influence on the perceived length of rods dropped on a floor (see also Wagman, 2003b). The influence of both haptic and visual feedback was checked in addition to an examination of the influence of practice without feedback. Three different lengths of rods (30, 60, and 90 cm) each of three different diameters (0.32, 0.95, and 1.59 cm) were dropped on a floor for subjects. Subjects were asked to report their lengths by positioning a moveable surface at a distance from the fixed edge of a desk so that the distance represented the length that would just be reached if the rod were extended from the edge of the desk. The same dropping and reporting apparatus that had been used by Carello et al. (1998) was employed for this investigation. Subject responses were found to improve in their consistency with practice, but their length estimation accuracies only improved when they were given visual feedback. With visual feedback, length estimates became increasingly correlated with actual length and decreasingly correlated with radius. Haptic feedback did not result in any improvement in accuracy, nor did additional practice.

Wagman, Hopkins & Minarik (2005a) presented results of an experiment in which rods of varying lengths and varying moments of inertia were dropped on a floor, and subjects were asked to estimate their lengths (see also Wagman, Hopkins & Minarik, 2005b; Carello et al., 2005). The experiments were done in an attempt to validate a hypothesis originally made by Carello et al. (1998) that the principal moment of inertia constrains perceived length (see also Anderson et al., 1997). To test this hypothesis, the PVC rods that were dropped in this experiment were internally weighted to decouple the principal moment of inertia from its length.

PVC pipes of 30, 45, and 60 cm in length and 1.5, 2, 2.5, and 4 cm in diameter were filled with lead shot and placed inside other PVC pipes of 6 cm in diameter. A fifth pipe was also used which contained no inner pipe. Listeners, separated from the stimuli by a curtain, heard the rods dropped onto the floor next to them and were asked to estimate the lengths of the rods by positioning a moveable flag at distances that corresponded to the lengths of the rods. This response system is similar, but not identical, to that used in previous research (as reported in Anderson et al., 1997; Carello et al., 1998; Wagman, 2003a). Surprisingly, actual length failed to account for any of the variance, and the principal moment of inertia accounted for only 33% of the variance in the results of the listening test. However, when rods of different lengths were analyzed separately, the principal moment of inertia accounted for much larger amounts of the variance (82-98%) in perceived length. The experimenter suggested two possible explanations for the puzzling results. The variety of stimuli tested may have been too small to demonstrate the desired phenomena, or alternatively, principal moment of inertia may not be the variable that governs length perception. The author suggested, “It may be that a ‘higher’ higher-order stimulation variable is specific to auditory perception of length.”

Kunkler-Peck & Turvey (2000) expanded on previous investigations of whether or not size is perceivable from objects involved in impact sounds, this time, using two-dimensional plates. The length along two dimensions, and thereby the shape, of rectangular plates was estimated by subjects who listened to the live presentation of the plates being struck at the centers by a pendulum. The plates were composed of steel, wood, and Plexiglas. It was found that subjects were good at estimating the relative dimensions of the plates, but the absolute sizes were generally underestimated. It is a possibility that the tendency to underestimate the absolute sizes may have been a bias caused by the reporting apparatus. The material influenced the absolute judgments made, but the relative dimensions of the judgments were not influenced by material. In further experiments, subjects were found to be relatively good at discerning between circular, triangular, and rectangular plates, even when the plates were presented in steel, wood, and Plexiglas. Additionally, subjects were nearly perfect at identifying the material of the struck plates, with only one of nine test subjects making one mistake. In addition to the interesting findings concerning the information found to be perceivable by listeners, the results also indicate that listeners were able to extract invariant information from signals of which multiple dimensions changed simultaneously.

Tucker (2003), in his Ph.D. thesis, investigated the perception of size, shape, and material of struck plates both underwater and in air (see also Tucker & Brown, 2002a,b,c, 2003). The intended application for such work is in the analysis of transient sonar signals. Square, circular, and triangular (equilateral) plates, all of the same surface area, were constructed from metal, wood, and plastic. In order to test size perception, three different sizes of square plates were constructed from the same three materials. The edge length of the medium-sized plate was double that of the small plate, and the edge length of the large plate was double that of the medium-sized plate. Monophonic recordings were made of all plates being struck at their centers by a metal spike, and these were presented to test subjects over headphones (diotically). Shape and material judgments were made using a three-alternative forced choice procedure, while size judgments were made in pairs using a visual response method in which the relative sizes of two squares were adjusted until the subject believed they corresponded to the relative sizes of the stimuli. Subjects could listen to the stimuli as many times as they wished.

In general, judgments based on underwater recordings were worse than those made based on in-air recordings. Material recognition was found to be good in both air and water, with the majority of confusions occurring between plastic and wood. However, some subjects consistently confused the two. Shape recognition was generally poor in both underwater and in-air conditions. Discrimination between gross size differences was found to be good, but subjects generally underestimated the differences between the sizes of the plates. Performance was best with metal plates and worst with plastic plates.

A model of the results was created which attempted to account for listeners predictions of material. The model, which first filtered the impact sounds with a Gammatone filterbank and then attempted to estimate the decay time of each channel, was able to account for the experimental results reasonably well. However, the model was unable to account for some tendencies, for example, that listeners were better able to judge the material of small plates.

Klatzky, Pai & Krotkov (2000) asked subjects to rate the similarity of synthesized sounds with respect to material and length. Five fundamental frequencies representing different lengths of bars were combined with different decay rates. Their rudimentary signals were presented to subjects in pairs, and in one experiment, the subjects were asked to rate the likelihood that the sounds came from the same material. The subjects were specifically told to ignore the frequency of the signal, and simply think about what material might produce it. Not surprisingly, in their multidimensional scaling analysis that was forced into two dimensions (McAdams et al., 2004), decay length was found to play a more important part in judgments of material similarity. In another experiment, new subjects were asked about the similarity between the lengths of the pairs of stimuli. In this experiment however, the subjects were told specifically to ignore how long the sound lasted, and to concentrate on what length of bar might produce it. Not surprisingly, decay rate did not play a role in length categorizations. In a final experiment, new subjects were asked to classify the synthetic sounds as glass, wood, steel, or rubber. They found that both decay rate and frequency influenced subject decisions. The authors concluded that simple, one-parameter models of material can be used for tasks when listeners must classify materials. However, it is the present author's opinion that tests that are more ecologically valid should be conducted before making such conclusions. The formulation of the task in many cases, and the description of the task to the subjects, appears to bias the results.

Lutfi (2001) performed an experiment similar to Lutfi & Oh (1997) in which the sounds of struck rods[4] were synthesized and presented monaurally to test subjects over headphones. The test subjects were asked whether the synthetic sounds of iron, aluminum, and wood rods were from solid or hollow rods. The synthesized materials and lengths of the rods were varied, and the subjects were told the material of each stimulus. The signals were not allowed to decay to inaudibility but were truncated after 1 s. This was at least a longer period than that used by Lutfi & Oh (1997).

In order to aid in his search for acoustic correlates to perceived hollowness, the author chose the stimuli so that the average performance accuracies would lie between 70% and 90% correct. Listeners were trained and were given feedback after every trial. It was found that some subjects performed the task according to the most appropriate acoustic relations, while others naively based their judgments on frequency alone. Some listeners even changed strategies during the test or for particular materials. The optimal strategy involved simultaneously listening to the decay rate, frequency, and intensity of the signal. By design, no one parameter alone could insure error-free performance; however, proper evaluation of all three parameters simultaneously would do so. The author suggested that limited sensory resolution was responsible for less than perfect performance even for the listeners who adopted the optimal decision strategy. In his discussion, the author additionally acknowledges that real-world acoustic events are very complex, and that it is almost certain there could be more cues available to listeners.

McAdams et al. (2004) asked subjects to rate the dissimilarity of synthesized versions of impacted bars. The authors stated that an aim of synthetic signals produced by physical models, like the one they employed, is that they resemble the real-world signals, which they are trying to mimic, as closely as possible. They continued by stating that such models can be very complex, particularly for an appropriate energy loss model, and then concluded that it was necessary to start with a simple model. Therefore, they wanted to validate the model described in the paper, as originally developed by Doutaut, Matignon & Chaigne (1998) for the

---

[4] The author referred to them as "bars".

simulation of mallet percussion instruments.

Both constant cross-section rectangular bars and bars with varying cross-section (like those of a xylophone and marimba) were used as stimuli in two experiments. The constant cross-section bars were synthesized using a physical model to have varying mass densities and different damping coefficients. Additionally, the bars with non-constant cross-section were synthesized to be of different lengths and to have different damping coefficients. Listeners made dissimilarity ratings, according to any criteria that they desired, on all possible pairs of sounds created by the physical model. The ratings were made by moving a cursor along a scale between "very similar" and "very dissimilar". The listeners could replay the stimuli as many times as they wanted. Stimuli covering a range of materials, from glass to wood, were synthesized, but metal was not synthesized because the model was not capable of representing the complex laws of damping as a function of frequency that are inherent in metals.

A multidimensional scaling analysis resulted in a solution of two dimensions for the results. This was the same number of dimensions as that of the physical model. Listening to the stimuli led the authors to conclude that one dimension resulted from pitch variation and the other from timbre variation, but that the timbre variation appeared "to have a temporal component and a spectral component." The latter dimension was identified as being related to the damping factor. The authors found that the decay rate was the dominant dimension. The authors were able to explain a significant amount of the variance in the listening test results by a linear combination of the spectral centroid and decay rate. They conclude that their experiment is one of the first to be able to use a combination of temporal and spectral descriptors to explain a perceptual result. This conclusion was in agreement with that of Klatzky et al. (2000), but in reaching their conclusion, a less rudimentary model has been used for the sound synthesis, so its value is theoretically greater. The results were, however, in disagreement with Lutfi & Oh (1997), who found that frequency was more important than decay rate and intensity. McAdams et al. (2004) suggest that this could be because the signals used by Lutfi & Oh (1997) were selected such that the differences between the sounds would be near the just noticeable difference thresholds of the listeners.

Lutfi, Oh, Storm & Alexander (2005) investigated the abilities of a large number of listeners to discern between recordings of impact sounds and synthetic impact sounds designed to match the recordings. Monophonic recordings of the stimuli being struck with hammers were made in a sound-treated room. The stimuli included a short pipe, a ceramic plate, an aluminum tube, a slab of wood, a juice glass, a large iron pipe, a ceramic bowl, a strip of metal, a metal chime, and a brass rod. Synthetic stimuli were generated with a rudimentary implementation of a “physically informed” model, like that suggested by Gaver (1993b). The model’s parameters were adjusted for each synthetic stimulus in an attempt to match the recordings. The results showed that only some of the listeners (approximately 26%) were able to reliably discern between the recorded and synthesized signals. Their work indicated that it was possible to generate sounds that could not always be discerned from monophonic recordings.

In a similarity-rating test, van den Doel, Pai, Adam, Kortchmar & Pichora-Fuller (2002) described perceptual experiments designed to test the validity of synthesized impact sounds. Two impact sounds were chosen for investigation: 1) a metal hammer striking a metal bowl and 2) a metal hammer striking a ceramic bowl. Monophonic recordings were made at a sampling rate of 22.05 kHz. The synthetic stimuli were synthesized using a progressive modal resonance model with which the number of modes can be increased at the cost of computational complexity. Test subjects were asked to rate the similarity of synthesized sounds and the original sound recordings. It was found that synthesized impacts could be created that were rated as perfectly similar to the recordings of the original impacts. The results suggest that the models could potentially be used in applications where it is necessary to generate real-time sound effects. However, it is not clear that they would be adequate for experiments seeking to understand human auditory perception. More experiments would be necessary to investigate this.

Giordano (2005) described experiments using both real-recorded and synthetic impact sounds. One of the experiments described in the thesis is a set of listening tests in which subjects were asked to rate and distinguish changes between the hardnesses of (hypothetical) hammers and the hardnesses of (hypothetical) objects struck with those hammers (see also Giordano, 2003; Giordano & Petrini, 2003). The experiment used synthetic stimuli. Three parameters of an impact sound model were varied to simulate changes in the hammer and the sounding object. While these experiments may be more of a test of the model than a test of perceptual abilities of listeners, it was found that subjects were, to a limited extent, able to discern between changes in features of the hammer and that of the object being struck. The properties of the struck object also influenced perception of hammer hardness and vice versa. This result is in contrast to that of Freed (1990), who found subjects were able to rate hammer hardness independently of the struck object. Giordano (2005) suggested that the fact that Freed (1990) specifically told his subjects to ignore the impacted object may have been responsible for the differences. However, it should also be kept in mind that the experiments of Freed (1990) were not done from synthesized stimuli, but from recordings of real impact sound events. This may also explain the differences.

In an additional experiment described by Giordano (2005), test subjects made dissimilarity ratings of monophonic recordings of struck plates. Subjects were free to rate the dissimilarity of pairs of stimuli based on any criteria that they decided on. The results showed that judgments made by most subjects were based on the properties of the struck object and not on the hammer used in the strike. Properties of the interaction between the hammer and plate were only relevant for some subjects. In general, the author concluded that the results of both of these experiments suggest that properties of the sounding object are those that provide the most information to listeners, and that the properties of the non-sounding object provide little information to listeners. This conclusion does not support the conclusion of Freed (1990), as described on page 28.

Giordano & McAdams (2006) described further impact sound experiments in order to investigate material perception (see also Giordano, 2003, 2005). Monophonic recordings were made of loosely supported glass, metal, plastic (Plexiglas), and wood plates being struck with a steel pendulum. The plates were square, 2 mm thick, and had areas from 75 to 1200 cm^2, corresponding to edge lengths between 8.7 and 34.6 cm. The stimuli were presented via headphones to subjects who sat in a soundproof booth. The test subjects could listen to each stimulus as many times as desired before deciding of which material the struck plate was made.

The results indicated that material identification performance was nearly perfect for gross material groups. Wood and plastic were assumed to form one of the categories, and metal and glass the other. The size of the plates did not influence gross material judgments. These results were consistent with Gaver (1988), Kunkler-Peck & Turvey (2000), Giordano (2003), and Tucker & Brown (2003). Signal duration, an issue ignored by Lutfi & Oh (1997), was found to explain the results of gross discriminations just as well as damping. The authors conclude that little support for the relevance of damping existed in their results.

Though gross material category distinctions were excellent, material judgments within gross material categories (wood-plastic, metal-glass), were poor. Material did not have a statistically significant influence in these within-category judgments. Plate area alone accounted for the results. Lutfi (2001) and Giordano (2003) had similar results, but in an experiment in which stimuli were presented live, Kunkler-Peck & Turvey (2000) previously found that subjects were also nearly perfect at judgments between wood and plastic.

An acoustic analysis helped to understand which factor test subjects used in making their erroneous judgments involving metal and glass stimuli. Signal frequency was found to correlate highly with the judgments. Frequency was also found to be a likely candidate for explaining judgments between wood and plastic but by itself could not completely account for the perceptual data. Other parameters such as the average signal loudness were needed to be used in combination with frequency to explain the judgments. The result that frequency is an important parameter in judgments within material categories is similar to the conclusions made by Klatzky et al. (2000) and McAdams et al. (2004), but in contrast to those of Roussarie (1999) who used synthetic stimuli in a perceptual investigation.

Ishibashi & Preis (2005b) described the results of experiments in which listeners were asked to compare the weights of iron balls dropped onto a floor. The goal of the experiments was to find just noticeable differences (JND's) for mass. Just noticeable differences have historically been measured for acoustical properties like frequency, intensity, and duration (e.g., Moore, 2003), but not for properties of the sources themselves. Binaural recordings of seven iron balls dropped from two different heights (45 and 85 cm) were made with an acoustic dummy head. The balls ranged in mass from 505 to 1515 g, in steps of approximately 200 g. Additionally, the balls dropped from a height of 45 cm were dropped onto both "hard carpet" ("hard" condition) and onto a blanket on top of the carpet ("soft" condition). The balls dropped from a height of 85 cm were only dropped onto the soft surface. By varying the drop height and surface onto which the balls were dropped, the energy of the impacts and the excitation and damping of harmonics was varied. Seven listeners were presented with pairs of recordings over equalized headphones. The listeners judged which of the two stimuli was the recording of the heavier ball.

From the results, mean just noticeable difference thresholds for mass were calculated. The results were somewhat dependent on the height from which the balls were dropped and the surface onto which the balls were dropped. Significant differences were found between balls dropped from 45 cm onto hard (JND: ˜200 g) and soft (JND: ˜275 g) surfaces, and between balls dropped onto the soft surface from 45 cm (JND: ˜275 g) and 85 cm (JND: ˜175 cm). No significant difference in the mass JND values was found between the balls dropped from 45 cm onto the hard surface (JND: ˜200 g) and the balls dropped from 85 cm onto the soft surface (JND: ˜175 cm).

Additionally, mean sharpness values (Zwicker & Fastl, 1999) and A-weighted sound exposure levels (L_{AE}) were calculated for each of the stimuli. L_{AE} was found to increase monotonically as the mass of the ball increased. The highest L_{AE} values were found for the balls dropped from 85 cm onto the soft surface, and the lowest L_{AE} values were found, as expected, for the condition when the balls were dropped from 45 cm onto the soft surface. In general, but not for every ball, the mean sharpness values were similar – the highest sharpness values occurred when the balls were dropped from a height of 85 cm onto the soft surface, and the lowest mean sharpness values were found for the case when the balls were dropped from

45 cm onto the soft surface. The floor surface, as opposed to the drop height, was found to effect the mean sharpness values the most. When the sharpness values of each of the different balls were compared to one another for each of the conditions, the balls that were dropped from 45 cm onto the hard surface produced the least variation in mean sharpness versus mass.

Both L_{AE} and mean sharpness were found to correlate strongly with subjective judgments of the balls' masses. The largest JND for mass was found for the condition with the lowest L_{AE} values – balls dropped from 45 cm onto the soft surface. This finding resembles the tendency for just noticeable intensity differences of pure tones to improve at higher intensities (Moore, 2003; Zwicker & Fastl, 1999). These results provided evidence that listeners were able to hear the masses of the balls dropped in this experiment. The findings indicate that the mass of balls is more difficult to hear when the balls are dropped onto soft surfaces, and that mass is also more difficult to hear when dropped from lower heights.

Ishibashi & Preis (2005a) described the results of further experiments in which listeners were asked to make further comparisons of the weights of balls dropped onto a hard carpet. Three test subjects listened to the recordings of the balls being dropped from 45 cm onto hard carpet, but which had now been normalized so that the maximum 2 ms Zwicker loudness values (Zwicker & Fastl, 1999) of each stimulus was set to one of two particular values (21 and 41 sone). The authors demonstrated that when the signals were normalized to 41 sone, the maximum Aures' sharpness values for each ball were nearly identical to their values before the normalization. In the condition of the signals being normalized to a maximum loudness of 21 sone, the maximum sharpness values changed somewhat dramatically and depended on ball mass. The range of maximum sharpness levels was greater in this condition (normalized to 21 sone) than in the unmodified condition.

For each loudness condition, listeners again compared the stimuli to one another. In these two conditions, maximum loudness was no longer available as a cue for listeners to perform the task. The results indicated that listeners were unable to perform the task either for the signals that had been normalized to 21 sone or for those that had been normalized to 41 sone. Correlations of the perceptual judgments and the sharpness were very poor, while they had previously been found to be very high before normalization. In summary, eliminating the maximum loudness variation between the signals dramatically decreased the performance of the subjects.

In the above experiments, the authors have employed two classical psychoacoustic measures of sensations, which may or may not be reasonable for characterizing the signals used in the tests: loudness and sharpness measurements. It seems that an ecologically valid result would be one in which the perceived characteristics of the dropped balls were found to correlate directly with physical properties of the sounds, and not with measures designed to estimate sensations produced by the sounds, measures which may or may not have ecological relevance.

Grassi (2002) described experiments in which listeners were asked to estimate the size of pine balls dropped on circular baked-clay plates[5]. Monophonic recordings of the sound events were manipulated in various ways and presented to test subjects over headphones. Subjects estimated the sizes of five 10 to 25 mm balls from the recordings in a four-alternative forced choice task. The subjects completed the test in five conditions, by listening to five variations of the recordings: 1) unmodified, 2) stimuli in which the bounces had been removed, 3) stimuli normalized according to their RMS power, 4) low-pass filtered, and 5) high-pass filtered. Subjects performed best when listening to the unmodified sounds. Performance in all but the high-pass condition was found to be significantly worse, and performance in the high-pass condition was close to being statistically significantly worse ($p = 0.06$). Performance in the condition in which the signals had been normalized to equate their RMS power levels was found to be worst. The results suggested that level may be one of many important cues to size perception.

Grassi (2005) described the results of live experiments in which listeners were asked to judge the sizes of seven pine balls dropped on circular baked-clay plates of three different sizes (see also Grassi, 2003; Grassi & Burro, 2003). The balls ranged from 10 to 50 mm in diameter and the plates were 165, 185, and 215 mm in diameter. The plates rested on a large, 40 mm thick foam block. In these experiments, most of the sound created in the impact was of the plate itself, not of the ball of which the size estimate was being requested. Therefore, the task was that listeners were asked to judge a property of the non-sounding object. The results could provide information on whether or not the sounding object carried information about the ball that impacted it. Subjects were told only that they would hear an object dropped on a plate, but were told nothing about the shapes or materials. They responded to the live stimuli by using a computer to draw a disk that corresponded to the perceived size of the ball.

All subjects correctly identified the shape of the balls as spheres, but the subjects' reports of the materials of the balls and plates were not always correct. The subjects also had trouble estimating the number of plates being used in the experiments. Subjects' size estimates were generally fairly accurate. Average ball size estimates were appropriately ranked with respect to the actual sizes, and absolute

---

[5] The author describes the plates as being made from "crate" in this article, but the plates are said to be made of "baked clay" in another article (Grassi & Burro, 2003).

estimates were reasonably accurate as well. However, in contradiction to Grassi (2002), plate size was found to affect the perception of ball size. Balls dropped on large plates were judged larger than balls dropped on small plates. Overall, ball size was somewhat underestimated, more so for small balls. However, the largest ball size (50 mm) was also underestimated more so than the second largest ball (40 mm). A power function with an exponent of approximately 1.5 was found to describe the relationship between judged sizes and actual sizes. Grassi (2003) had previously shown similar results for visual estimates of ball size using the same response method. Ball size was generally underestimated in the visual test, but to a lesser degree than in the auditory task.

Although the effect of the plate size was apparently not nearly as large as the effect of the ball size in the auditory experiments, the author concluded that listeners were not able to hear independent features of the objects involved in the impact events. Freed (1990), as well as Grassi (2002) himself, previously found that listeners *were* able to separate the properties of the non-sounding object from the sounding object. Grassi (2005) attributed the differences between his study and that of Freed (1990) to the fact that the subjects in the experiment of Freed (1990) were given more information about the experiment prior to the experiment being conducted.

An acoustical analysis suggested that the centroids of the spectra of the sounds could indicate the sizes of the balls. Differences in the plates would affect this centroid but should mainly have the effect of shifting all frequency components up or down, depending on whether the plate size was decreased or increased. Average RMS power was found to be an acoustic parameter that was good at predicting subject performance, but none of the acoustic parameters analyzed (average RMS power, peak amplitude, spectral centroid, event duration, and bounces) could explain all of the data. Actual ball size was the best predictor of subject performance. These results suggest that more complex acoustic descriptors are necessary to understand how information concerning size was carried to the test subjects in this experiment.

Summary of Impact Sounds A condensed summary of the impact sound investigations described in this subsection is presented in Tables 2.1 and 2.2. Each row in the tables lists the authors and date of publication of the work, the event tested, the properties of that event that were investigated, and the stimulus presentation method used in the perceptual experiments. Additionally, the page numbers on which the experiments are discussed in this thesis are provided in the rightmost column. Table 2.1 lists research published before the year 2000, and Table 2.2 lists research published from 2000 to present.

Table 2.1: Summary of impact sound research prior to the year 2000 that has been reviewed in this subsection. Entries are sorted by year of publication. The column describing *Presentation* (Method) indicates the way(s) in which the stimuli were produced for the test subjects. *Mono* indicates that monophonic recordings were made of the stimuli and reproduced for the test subjects, *Binaural* indicates that binaural recordings were used, *Synthetic* indicates that stimuli were synthesized (always monophonically), and *Live* indicates that stimuli were presented live. A page number is listed on which the reader may find more information in this thesis.

Publication Authors (Year)	Event Tested	Property Questioned	Presentation	Page
Warren, Jr. & Verbrugge (1984)	Dropped Glass	Breaking/Bouncing	Mono	23
Repp (1987)	Clapping Hands	Identity, Gender, Config.	Mono	24
Warren, Jr. et al. (1987)	Bounced Balls	Bounciness	Live	24
Gaver (1988)	Struck Bars	Material, Size	Mono	25
Wildes & Richards (1988)	Struck Objects	Material	-	24
Freed (1990)	Struck Pans	Hardness	Mono	28
Li et al. (1991)	Walking Sounds	Gender	Mono	28
Heine et al. (1993c)	Dropped Balls	Number	Binaural	29
Anderson et al. (1997)	Dropped Rods	Length	Live	30
Lakatos et al. (1997)	Struck Bars	Shape, Material	AB Stereo	27
Lutfi & Oh (1997)	Struck Bars	Material	Synthetic	25
Pittenger et al. (1997)	Shook/Stirred Balls	Size	Live	29
Carello et al. (1998)	Dropped Rods	Length	Live	30
Houix et al. (1999)	Struck Bars	Similarity of Bars	Mono	27
Pittenger & Mincy (1999)	Shook/Stirred Granules	Size	Live	30

Table 2.2: Summary of impact sound research from the year 2000 to present that has been reviewed in this subsection. Entries are sorted by year of publication. The column describing *Presentation* (Method) indicates the way(s) in which the stimuli were produced for the test subjects. *Mono* indicates that monophonic recordings were made of the stimuli and reproduced for the test subjects, *Binaural* indicates that binaural recordings were used, *Synthetic* indicates that stimuli were synthesized (always monophonically), and *Live* indicates that stimuli were presented live. A page number is listed on which the reader may find more information in this thesis.

Publication Authors (Year)	Event Tested	Property Questioned	Presentation	Page
Klatzky et al. (2000)	Struck Bars	Material/Length Similarity	Synthetic	34
Kunkler-Peck & Turvey (2000)	Struck Plates	Size, Shape, Material	Live	33
Cooper et al. (2001)	Struck Rods	Length	Live	31
Lutfi (2001)	Struck Rods	Solidity/Hollowness	Synthetic	35
Grassi (2002)	Dropped Balls	Size	Mono	41
van den Doel et al. (2002)	Struck Bowls	Similarity	Mono/Synthetic	37
Tucker (2003)	Struck Plates	Size, Shape, Material	Mono	33
Wagman (2003a)	Dropped Rods	Length	Live	32
McAdams et al. (2004)	Struck Bars	Similarity	Synthetic	35
Giordano (2005)	Struck Objects/Plates	Hardness, Similarity	Mono/Synthetic	37
Grassi (2005)	Dropped Balls	Size	Live	41
Ishibashi & Preis (2005a,b)	Dropped Balls	Mass	Binaural	39
Lutfi et al. (2005)	Struck Objects	Realism	Mono/Synthetic	36
Wagman et al. (2005a)	Dropped Rods	Length	Live	32
Giordano & McAdams (2006)	Struck Plates	Material	Mono	38

2.2.3 Protocol Studies

A number of researchers have performed studies on large sets of everyday sounds that cover a wide range of sound events, but only a few will be discussed in detail. Other studies not discussed below have examined the ability to recognize sounds, the ability to recognize auditory scenes (Peltonen, Eronen, Parviainen & Klapuri, 2001), recognition memory (Miller & Tanis, 1971), the effects of context on identifiability (Ballas & Mullins, 1991; Spanik & Pichora-Fuller, 1999), perceptual similarity (Bonebright, 2001), and onomatopoeic representations of sounds (Takada, Tanaka & Iwamiya, 2006). They have also offered advice on where to find and how to select sounds for protocol studies (Shafiro & Gygi, 2004). Attempts to build algorithms that can computationally recognize environmental sounds have also been attempted (e.g., Goldhor, 1993).

VanDerveer (1979) was among the first to investigate auditory perception using the ecological approach to perception. She performed experiments with both children and adults in which she produced sounds for them and asked them to describe what they heard. She asked the question in such a way that the subjects were not led to believe that she was asking for descriptions of the sensations produced by the sounds nor that she was asking for the subjects to identity the sound producing events, but subjects were found nearly always to describe the sounds in terms of the source events. Only when subjects were unable to identify the source did they revert to more abstract perceptual descriptions. The most frequent confusions were found between stimuli having similar temporal structure. She proposed that listeners are likely to be more capable of identifying objects involved in sound events when those objects are involved in several kinds of events (e.g., impact and scraping events), as opposed to repeated involvement in the same kind of event (e.g., only impact events).

After initial experiments showed that adults were able to identify many sounds from monophonic recordings, experiments were conducted in which preschool aged children (4–5 years old) were asked to do the same. Two groups of children performed the test, one group listening to stimuli presented live, and another listening to monophonic recordings presented over a loudspeaker. Performance was worse for the presentation of recorded stimuli, but it was unclear if this was simply due to the presentation of recorded stimuli being less interesting for the children, the fidelity of the recordings, or some other factors. In general, the children were often able to identify the sounds, but not as well as the adults were. Additionally, the mistakes that were made covered a much greater variety of sound events.

In an additional experiment, VanDerveer performed tests to investigate memory of everyday sounds. She asked subjects to listen to a set of sounds, listen to another set of sounds ten minutes later, and listen to another set of sounds three months later. The task of the subjects was to indicate whether each of the sounds that they heard ten minutes and three months later were in the original set of sounds. There was little difference between the results from ten minutes later and from three months later. In general, it was more likely that subjects claimed they had heard a sound before when they actually had not, than for the subject to claim they had not heard a sound before when they actually had. The author pointed out that this is similar to how it is less likely that someone would *not* recognize a known face than it would be for that person to mistakenly recognize an unknown face. Additionally, she discussed the possibility that memory for recognition of sounds may depend on verbal labeling of those sounds that a subject may do in his or her head. She proposed that if labels are used to recognize sounds, then people should not be able to differentiate between sounds that have the same name but sound different.

In a final experiment, VanDerveer (1979) asked subjects to perform a free classification (sorting) task and a paired-comparison task in which subjects were to group and compare the acoustic similarity of sounds. The results were somewhat difficult to interpret, but the author found that synthetic sounds tended to be outliers. The most important acoustical attributes seemed to be classifiable as temporal and spectral parameters. Sounds with similar temporal attributes such as percussive or continuous sounds, sounds with rhythmic patterning, and sounds with similar attack or decay times were often grouped together. Sounds with similar spectral characteristics, such as the material, texture, and sound when undergoing changes in form (e.g., rattling or bending paper), were also noted to form a basis for categorizations. From the results of this experiment, it was hypothesized that the frequency content may be predominately responsible for carrying information concerning surfaces and substances, while temporal information may carry information about the action taking place in the events. It was pointed out that although the subjects were asked to compare and sort the sounds based on acoustical characteristics, there is no guarantee that they did this. The author suggested that future work was necessary to disentangle these issues.

Gaver (1988) described identification experiments in a part of his Ph.D. thesis (see a description of additional parts of his thesis on page 25). In contrast to VanDerveer (1979), after initially asking subjects, “What do you hear?”, the subjects were probed to supply more details concerning what they heard. Subject performance was good. The main finding of interest worth mentioning here is that he found subjects tended to over specify what they had heard. In other words, they tended to describe a more specific event than actually occurred. He gave the example that instead of stating that a hollow metal object had been struck, a subject might suggest that a tin can fell over. He suggested that this could have been due to the way the subjects were questioned, subjects' inabilities to describe the events abstractly, or possibly because of real tendencies to try to accurately identify events.

Ballas (1993) performed a number of experiments investigating acoustic, ecological, perceptual, and cognitive factors that were common to the identification of a set of 41 environmental sounds. The author measured the identification time, accuracy, frequency of occurrence, spectral and temporal properties, as well as other parameters. Frequency of occurrence was measured by asking test participants to write down the first everyday sound heard after a small alarm altered them to do so at various times during a day. The author found that 75% of the variance for stimulus identification time could be attributed to the acoustic properties of the stimulus and the frequency with which it occurs outside the laboratory. In general, a great amount of data was collected from which it was concluded that there were many acoustic and cognitive factors, which should be taken into account in the development of theories concerned with the identification of everyday sounds.

Fabiani, Kazmerski, Cycowicz & Friedman (1996) developed naming norms for a set of 100 sounds to be used in sound identification tasks. The sounds, all trimmed to be 400 ms in length, were made available for download from a web site. Tests in which the subjects attempted to identify the sounds were performed by 77 young adults, 41 older adults, 61 children of various age groups, and 17 subjects that were suspected to be suffering from Alzheimer's disease. All subjects, including the elderly, who had ages between 61 and 88 years old, were required to have audiograms indicating less than 40 dB HL of hearing loss. For the children, sound-naming performance increased as age increased. For the remaining subjects, sound-naming performance decreased with normal and pathological aging. Sound and picture naming performance were found to be correlated. The authors proposed that the resulting normative data could be used to guide the selection of sounds to be used in future sound identification tasks.

Marcell, Borella, Greene, Kerr & Rogers (2000) developed a set of 120 sounds to represent a variety of acoustic events – similar to that of Ballas (1993) and Fabiani et al. (1996), but with longer sounds. The authors felt that the brief duration of the sounds developed in previous studies may limit their usefulness. Instead of trimming the sounds to be a particular length, the authors set the length of each sound in such a way that sound event had sufficient time to occur naturally. The intended use of the sounds was for psychological tests to investigate information organization in memory and word-finding abilities, which may degrade with age or neurological impairments. A scoring system including acceptable responses for each of the sounds was developed, along with norms for response times, familiarity, pleasantness, and complexity. The sound files are freely available for research and clinical use.

Gygi, Kidd & Watson (2004) performed experiments to determine the influence of spectral and temporal factors on the identifiability of a set of 70 everyday sounds (see also Gygi, Kidd & Watson, 1999, 2000; Gygi, Watson & Kidd, 2000; Gygi, 2001). To investigate the contribution of spectral properties, the set of sounds were high-pass, low-pass, and band-pass filtered with varying cutoff frequencies. Subjects listened to the stimuli using headphones. Recognition performance was generally quite good under all conditions. For the band-pass sounds, best recognition accuracy occurred for sounds that had been filtered to exclude frequencies outside of 1200 and 2400 Hz. In general, the results were comparable to that of tests involving the filtering of speech, but with more information at high frequencies. The nature of the generation of speech and everyday sounds was cited for the likely explanation for the differences between speech and everyday sound perception. Similar to speech, the identifiability of everyday sounds is apparently quite robust.

Next, a vocoder processing technique was used to determine if the same set of everyday sounds could be recognized with limited spectral information. Vocoders of one and six channels were used. The results indicated that temporal information was in many cases sufficient to produce identification accuracies of at least 50% (chance level was 1.4% correct). As with speech, temporal structure is apparently also sufficient for the identification of many everyday sounds. The increase of spectral information by adding vocoder channels was also shown to enhance identifiability. Furthermore, it was found that subjects who had heard the original unprocessed sounds before performed much better than when they had not heard the original sounds previously.

Shafiro (2004b) performed listening tests in which normal-hearing listeners were asked to listen to 60 everyday sounds processed by a vocoder simulation of a cochlear implant (see also Shafiro, 2004a; Shafiro & Gilichinskaya, 2004). This was similar to the work of Gygi et al. (2004), but with a larger number of vocoder channels tested. It was found that increasing the number of channels of the vocoder to 32 resulted in improved identification accuracy for most sounds, but vocoders having 24 and 32 channels (the greatest number of channels tested) actually caused a decrease in performance for some sounds, compared to the 16-channel vocoder. It was suggested that this could have been because the filter delay for each channel was not constant, and therefore resulted in asynchronous spectral components. In general, the sounds processed by the vocoder were not recognized as accurately as the unprocessed signals. Time-varying sounds with narrow-band resonances were found to be the most difficult to recognize. The main conclusion of the thesis was that cochlear implants may require more spectral channels to convey information from environmental sounds than are required for speech perception. Speech intelligibility rates for single vowels may reach 90% with eight channels (e.g., Dorman, Loizou & Rainey, 1997), but eight channels was only sufficient to produce approximately 60% correct identification of environmental sounds.

Heller (2002a) described experiments in which listeners were asked to compare the realism of recordings of sounds events to recordings of sound *effects* like those that would be created by Foley artists (see also Heller, 2002b). Foley effects may be considered as exaggerations or caricatures of the sounds which they attempt to imitate (VanDerveer, 1979). In general, the real events were rated as more realistic, but some subjects consistently rated the Foley effects as more realistic. In an additional experiment, an attempt was made to synthesize sounds that would be perceived as being more realistic than the real events that they were designed to represent. Three events were chosen for this: walking in mud, walking in leaves, and crushing eggshells. This was done by extracting and recombining acoustic features of the stimuli that were thought to produce the action and the material. In two of the three cases, listeners rated the newly synthesized version as being more realistic than the real sounds or the Foley sounds. Apparently, for these two cases, the exaggerated characteristics of the Foley sounds were somehow convincing for listeners.

2.2.4 Information Common to Sound Events

There are some types of information commonly provided by many types of sound events. For example, most sounds provide information concerning the position – distance and direction – of the event relative to the listener. Changes in this distance and direction may also be provided. Additionally, information about the environment in which the sound event occurs may be provided through the reflections that have been structured by the environment. Therefore, information available to a listener need not originate from physical objects directly interacting with one another, but it may also be provided via reflections, diffraction, occlusion, and other acoustic phenomena.

For example, the size of a room may be perceivable when a loudspeaker generates sound in a room. Similarly, the distance to a wall may be perceivable because the wall itself structures sound produced by an active, sound-producing listener. This subsection will present a few examples of research concerning these types of phenomena.

Information on Occlusion

Gordon & Rosenblum (2004) asked blindfolded subjects whether they could fit through a doorway-like aperture on which recorded crowd noise was being reproduced by six loudspeakers on the other side (see also Gordon & Rosenblum, 2001). Russell (1997) had previously provided preliminary evidence that it is possible for listeners to determine whether or not there is an occluding surface between a loudspeaker and listener. An "acoustic shadow" is therefore potentially informative about the size and shape of the obstruction creating that shadow. In separate experiments, Gordon & Rosenblum (2004) varied the width and height of a doorway, and asked subjects whether they could fit through the doorway without turning their body or ducking their heads. Subjects stood in front of the doorway. Opening widths were adjusted between 5 and 95 cm, while the height of the aperture was adjusted between 122 and 188 cm. For the height experiments, the sound level was additionally adjusted to one of three levels and varied between trials.

The results indicated that subjects were moderately good at being able to differentiate between apertures through which they could fit and those through which they could not fit. The results from these auditory experiments were similar to that from visual experiments (Warren & Whang, 1987). Ratios of the perceived passable aperture sizes to subject heights and shoulder widths were calculated for all subjects. It was found that this ratio was not significantly different across subject size, providing evidence that subject judgments were made on body-referential dimensions.

Robart & Rosenblum (2005a) asked blindfolded listeners to judge the shapes of sound-obstructing objects. A square, circle, and triangle, each with a surface area of 7575 cm^2, were placed in front of an array of eight loudspeakers producing broadband noise. Listeners were allowed to move their heads before making their judgments. Though performance varied across subjects, subjects on average were found to be able to differentiate between the square, circle, and triangle at levels better than chance. Some listeners performed the task nearly perfectly. When the sound level of the sound produced by the loudspeakers was also varied between trials, subjects performed significantly more accurately than when a fixed level was used for all presentations. Apparently, listeners used information other than the absolute intensity produced by the loudspeakers for their judgments, and shape information was made even more perceptually accessible with the additional variation of level. This experiment illustrates an ecological listening point that extraction of source invariants is made easier by changes to irrelevant cues so that listeners may more easily hear what does not change (McAdams & Bigand, 1993).

Information on Distance and Reachability

Rosenblum, Wuestefeld & Anderson (1996) asked blindfolded subjects to judge whether they could reach a live sound source in their environment (see also Wuestefeld & Rosenblum, 1993). The research is an excellent example of an ecologically motivated listening task in which both the stimuli and a response to the stimuli are ecologically relevant. In short, the idea suggests that the information concerning the environment as registered by a person depends on the person registering it (Gibson, 1966). Thus, what is reachable for one person may not be reachable for another. In the experiment, the sound source was a mechanical shaker in which popcorn kernels were placed inside. In one case, the subjects were asked to judge whether they could reach the source simply by extending their arm, and in another case, the listeners were asked if they could reach the source when allowed to bend at the hip while simply maintaining contact with the seat of the chair. Estimates of "reachability", measured as the distance at which 66% (2 out of 3 repetitions) of the trials received "yes" responses, were generally very accurate for both conditions, and for both short- and long-armed listeners. However, long-armed listeners performed better. When judgments were scaled to the subjects' respective body dimensions, judgments for both groups of subjects were similar. On average, errors in reachability estimates were 11 cm for the arm-only condition and 18 cm for the condition in which the subject could bend at the waist. The results indicated that distance perception in a natural environment is better than has previously been found in lab experiments. The results suggest that reachability performance based on auditory information alone may rival or exceed that of performance when the task is performed visually (Carello et al., 2005).

In a second experiment, some subjects completed a similar task as previously described, and another set of subjects completed the task with the sound reaching their right ear attenuated. The right ear of each subject in this group was blocked with both a foam earplug and an over-the-ear headphone, together with a combined attenuation of approximately 35-40 dB. Subjects from both groups completed two conditions: one in which they could not move their head[6] and another in which they were required to move their head. No significant differences were found between the conditions in which the listeners' heads were fixed and the condition in which the listeners were required to move their heads. This result, along with that of Ashmead, Davis & Northington (1995), in which it was found that distance perception was improved when listeners walked during stimulus presentation, suggest that large listener movement is required before a distance-perception advantage from movement can be observed. Furthermore, the results indicated

---

[6] Their heads were rigidly fixed by the use of a bite bar.

that subjects performed the task more consistently in the condition where full binaural information was available, but no more accurately.

Russell & Schuler (2001) further investigated reachability and found that if the target was at the ear-height of standing listeners, the subjects performed significantly better than if the stimulus was produced from waist-height. Furthermore, if the stimulus (a loudspeaker producing a duck call) was directed towards the head of the subject rather than simply directly at the waist of the subject, accuracy improved. It was therefore concluded that perceived reachability is influenced by the amount of direct sound compared to the amount of indirect (i.e., reflected) sound reaching the listener.

Information from Reflected Sounds

Kim, Stoffregen, Ito & Bardy (2005) demonstrated that blindfolded listeners could coordinate their body movement to that of a wall in a moving "room", simply by using information from the sound in the room (see also Ito, Stoffregen, Donohue & Nelson, 2001, for previous work using blind test subjects). A 2.4 m cube sat on wheels and could be moved back and forth by an electric motor. Listeners stood inside the cube and were asked to sway their bodies along with the motion of the front wall. They were asked to sway in such a way that a constant distance was maintained between their body and the wall. A magnetic tracking system was used to track the position of the subjects relative to the wall. Two conditions were tested by each subject. In the first condition, loudspeakers were mounted to the corners of the cube, facing inwards. In the second condition, the loudspeakers were mounted on poles attached to the floor, which did not move when the cube moved. In both conditions, subjects were found to successfully follow the movement of the wall. The results suggested that normal-hearing listeners could extract information from reflected sounds.

Robart & Rosenblum (2005b) described investigations to determine whether listeners could discern between rooms, based on binaural recordings of sounds produced in those rooms. Five different sounds, including human speech, a live cowbell, and three synthesized series of noise bursts, were recorded in four rooms. The four rooms were a bathroom, a classroom, a gymnasium, and a small laboratory. The results indicated that listeners were remarkably good at identifying the rooms correctly, with an average accuracy of 78%. Additionally, there was a significant influence of the stimulus on identification accuracy, but there was no obvious explanation as to why the task was easier with some stimuli.

Sandvad (1999) had previously performed similar experiments, but had found a wide variety of performance differences between subjects. Some could perform the task nearly perfectly, but others made many errors. Some subjects used only the reverberation time, while others were found to take advantage of the direct-to-reverberant energy ratio. The author also presented evidence that subjects listening to binaural recordings of a speech signal could make 70% accurate judgments concerning where in the room the recording was made. Subjects did this by pointing to photographs of the rooms.

Weisser (2004), as a part of his master's thesis, conducted listening tests in which test subjects were asked to identify pairs of binaural recordings that were made in the same rooms. The rooms included various small rooms such as a meeting room, a small department library, a classroom, a radio station talk studio, and others. The rooms ranged in volume from 26 to 190 m^3; four were approximately 100 m^3. In each room, music and speech samples were played from a loudspeaker and were recorded by an acoustic manikin. Multiple recordings were made with the loudspeaker and acoustic manikin in various positions in the rooms. The task of each subject in the listening test was to listen to four recordings on each trial and to identify the pairs that were recorded in the same room. For a given trial, all four recordings were of the same program material, but the recordings making up the pairs were made with the loudspeaker and acoustic manikin in different positions.

The task was therefore one in which listeners were asked to recognize rooms from the influence of the room on the program material played in the room. Success at the task required that subjects could differentiate between sound qualities related to the room itself and sound qualities related to the positions of the loudspeaker and manikin (e.g., spatial location, level, direct-to-reflected energy ratio). All nineteen subjects were able to perform the task at greater than chance level, and some were perfect or nearly perfect. Ten pairs of rooms were found to be discernible at better than chance accuracy, and only two were not. The author found that a measure of low frequency reverberation time could explain many of the results, but room volume was a poor predictor of the results. Although the tasks were slightly different, the results appear to demonstrate an even more acute ability to hear the differences between rooms than was later demonstrated by Robart & Rosenblum (2005b). The test performed by Robart & Rosenblum asked listeners to identify recordings made in a bathroom, classroom, gymnasium, and a small laboratory. The rooms used by Weisser (2004) were of rooms that were much more similar to one another.

Information on Sound Source Motion and Facing Angle

Acoustic information affords perception of sound source motion (e.g., Shaw, McGowan & Turvey, 1991; Jenison, 1997). The ability to predict details concerning an impact, including its expected time, is an important skill for navigation in an environment (e.g., Lee, 1976; Neuhoff, 2001a). Listeners likely use cues such as higher-order, time-varying amplitude, Doppler shift, and interaural differences to predict the so-called "time-to-arrival" (Guski, 1992; Carello et al., 2005).

Rosenblum, Wuestefeld & Saldaña (1993) conducted listening tests in which subjects listened to truncated monophonic recordings of an approaching vehicle and were asked to indicate when the car would have reached the listening position if the recording had continued. The recordings were edited so that in some conditions, the subjects heard the recording up through a point just after the vehicle passed the listening position, and in other cases, the recorded sound was replaced with silence at specified points before the vehicle had passed the recording position. In addition to the variable point in time at which the recording was muted, the total time of which the recording was audible was also varied. Jenison (1997) had previously provided a physical analysis to demonstrate that information for position, velocity, and time-to-arrival is available in time-varying measures of interaural-time-delay, Doppler shift, and sound level. Some of this information requires binaural hearing to perceive, while some of it does not. For the experiments described by Rosenblum et al. (1993), it should be kept in mind that because monophonic recordings were used, binaural information was of course absent.

It was found that judgment errors increased for the conditions in which the audible portion of the vehicle approach was further separated from the actual time in which the vehicle passed the microphone, even when total audible-stimulus times were equated. The total duration of time in which the approaching vehicle was audible was not found to significantly influence the results. Unless they were given feedback, it was found that subjects were not better at estimating the arrival time even when they could hear the complete signal. Feedback resulted in improved performance.

Neuhoff (2001a) conducted multiple experiments, which showed that approaching sound sources were perceived to be closer to a listener than receding sound sources (see also Neuhoff, 1998). It was suggested that this bias was due to evolutionary adaptation that could aid in survival. In an initial experiment of a more classical psychoacoustic style, subjects listened to tones over headphones that were either rising or falling in intensity. These signals were intended to simulate approaching and receding sources, respectively. A change of 30 dB occurred over a period of 1.8 seconds, either rising or falling, between either 40 and 70 dB or 60 and

90 dB. In a paired comparison task, subjects were asked to compare the amount of loudness change between paired combinations of the stimuli – either with the first stimulus rising and the second stimulus falling or with the first falling and the second rising. The results indicated that rising sound levels were judged to change by a greater amount than falling sound levels. Additionally, the stimuli that rose from 60 to 90 dB were judged to change more than the stimuli that rose from 40 to 70 dB. The author pointed out that from a survival standpoint, higher intensity would indicate a closer source, and therefore less time to react.

Similar results were found in live experiments conducted outdoors. Blindfolded listeners heard a loudspeaker moving towards them or away from them along a cable supported by posts at either end. In one case, the loudspeaker started at the far end of the cable and moved to the midpoint. In a second case, the loudspeaker started at the midpoint and moved to the end of the cable closest to the listener. The opposite conditions were also presented, with the loudspeaker starting near the subject and moving to the midpoint of the cable, and with the loudspeaker starting near the midpoint and moving to the far end of the cable. On each trial, subjects were asked to judge the final position of the loudspeaker, which played either tonal or broadband noise during its motion. Two subject groups performed the test, one making verbal estimates of the starting and ending loudspeaker positions and the other by walking, blindfolded along the cable, to the starting and ending positions of the loudspeaker. The results indicated that in the conditions when the loudspeaker was approaching the listener, the final position was judged closer than in the condition in which the loudspeaker was receding from the listener. These results again suggested that the auditory system was acting in a cautious manner when the sound source was approaching, a defensive posture that could be useful for survival. Subjects were more accurate with broadband stimuli than with tonal sounds. Additionally, better accuracy was found when the subjects walked to the perceived stopping position.

Neuhoff (2003) asked listeners to judge the facing angle of a loudspeaker reproducing human speech (see also Neuhoff, 2001b,c). Listeners positioned an identical, silent loudspeaker at the angle at which they believed the active loudspeaker was facing. They performed this task in a condition where the loudspeaker was stationary during the stimulus presentation, but also for a condition in which the loudspeaker rotated during the stimulus presentation. In the latter case, subjects judged the final facing angle. Subjects performed better when the loudspeaker rotated while the stimulus was playing, and they performed best when the loudspeaker was rotated *toward* the listener during stimulus presentation. Average absolute error for the dynamic condition in which the loudspeaker was positioned approximately 1 m from the subject was approximately 35 ° For all conditions, subjects performed best when the loudspeaker was facing them, but in the static condition, there were many reversals between 0 ° (directly facing the subject) and 180 ° (facing directly away from the listener). In the dynamic condition, these reversals disappeared. The results underscore the importance of dynamic listening situations and the benefit listeners derive from them. The dynamic nature of the real world and that of the stimuli in the test have again made perception of invariants easier, not more difficult. These results have implications for any source that may have directional characteristics, including speech.

2.3 Experimental Methods

A variety of stimulus presentation methods and test response methods have been directly shown and indirectly suggested to influence the perception of everyday sounds. Some of these methods will be compared and discussed here.

2.3.1 Stimulus Presentation Methods

In the experiments described in this chapter, researchers have used a variety of methods to generate and present stimuli to test subjects. Among others, popular ways that stimuli can be produced are by

- synthesizing them using mathematical models,
- making monophonic recordings using a single microphone,
- making stereo recordings using classic stereo recording techniques such as AB stereo, XY stereo, and MS stereo,
- making binaural recordings using an acoustic manikin, or
- producing them live.

If not presented live, the presentation of stimuli can occur over headphones or loudspeakers. If headphones are used, the stimuli can be presented diotically or dichotically. Monophonic stimuli can be presented to a single ear or diotically to both ears. Stereo stimuli would typically be presented dichotically to both ears if the full information in the recordings is utilized. If loudspeaker presentation is used, the stimuli can be presented in a normal room or in an anechoic chamber, from a single loudspeaker or from multiple loudspeakers. There are of course many other ways that stimuli can be recorded and presented, but these are the common ways and the common decisions that need to be made for each parameter.

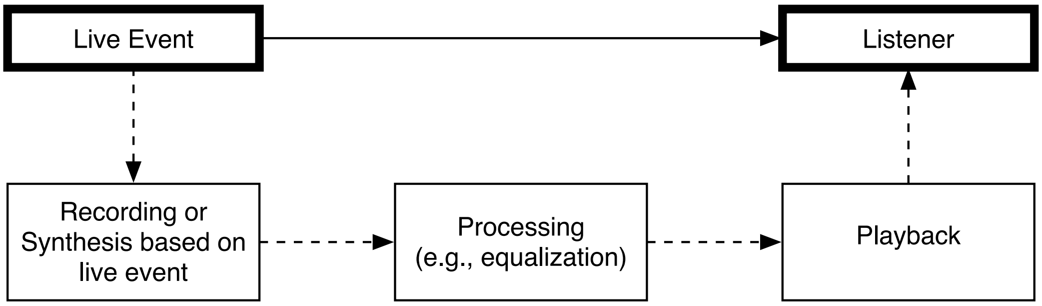


Figure 2.2: Illustration of possible ways in which sounds can be presented to listener. The live event may be presented to a listener directly, or a recorded or synthetic version of the event may be created, processed, and presented to a listener with the help of reproduction equipment.

Each has advantages and disadvantages. A diagram illustrating possible routes for presenting sounds is shown in Figure 2.2. It is apparent that with live presentation, where the sounds are presented directly to listeners, there are fewer stages at which perceptual degradations can be introduced.

Monophonic recordings do not maintain any of the spatial information conveyed by interaural time differences and interaural level differences that would have been available to the listener if he or she had been present at the time of the original sound event. Listening to monophonic recordings presented to both ears of a subject over headphones typically results in the listener perceiving the sound as occurring within his or her head. Loudspeaker presentation of monophonic recordings overcomes the problem of the sound being perceived inside the head of the subject, but recordings should be made in an anechoic chamber if they will later be played back in an acoustically normal room. Otherwise, the listener will hear a jumbled combination of the influence of two rooms. With this method, however, the directionality of the source and other spatial aspects of reflections that would occur in a normal room would still not be properly conveyed.

Common stereo recording techniques such as AB stereo and XY stereo allow the recorder some control over the size of the sound image that is ultimately perceived by a listener, but these techniques still suffer from the fact that the source seems to be playing inside of the head of the listener when he or she listens using headphones. When listening to stereo recordings with loudspeakers, problems still exist that are similar to those described above for monophonic recordings presented over a loudspeaker.

In theory, properly made binaural recordings made using an acoustic manikin with microphones at the entrance to its ear canals or at the position of the eardrum should be a very accurate representation of the sound reaching the ears of a listener in place of the manikin (Møller, 1992). These recordings would typically be played back to listeners over headphones, and if so, consideration must be made for whether the microphones in the head of the acoustic manikin were at the position of the eardrum or if they were at the entrance to the ear canal. If the recordings are made at the entrance to the ear canal, then the recordings should be played back from the entrance to the ear canal. If the recordings are made at the eardrum, then an equalization must be done to account for the fact that the recordings will not be played back from the position of the eardrum. This is necessary because the sound would effectively travel through two ear canals if the recordings were made at the position of the eardrum and then played back via headphones to the eardrums of the wearer, thereby distorting the signal. Furthermore, the influence of the headphones or insert earphones must be compensated for. Accurate signal presentation requires that an equalization should be performed in order to reduce further distortions being introduced by the imperfect frequency response of the playback system.

While the binaural recording technique can produce extremely realistic sound images, the technique does not accommodate for the listener moving his or her head. If the listener moves his or her head, the entire auditory environment will move with it. The auditory environment will not stay in a fixed position as the original physical environment would. Therefore, for most realistic reproduction, the listener should not move his or her head. Few researchers have used the binaural technique for everyday sound research (e.g., Heine et al., 1993c; Peltonen et al., 2001; Robart & Rosenblum, 2005b; Sandvad, 1999; Ishibashi & Preis, 2005a,b), but it offers excellent advantages over monophonic recordings for the study of everyday sounds.

Hybrid methods, such as convolving head-related transfer functions with monophonic recordings made in an anechoic environment, are also possible. Such a system can be implemented in which this process is done in real time, taking into consideration data from a head tracking system worn by the test subject. This particular technique can be used to provide the ability for the listener to move his or her head during the listening test while having the auditory environment stay in its proper position. The method, although technically difficult to implement, can be used so that the auditory "image" perceived by a listener does not move around in the physical environment as the subject moves his or her head. The intended physical position of the sound source in relation to the environment in which the subject is listening can be represented. However, the fact that the signal being convolved with the head-related transfer functions is only monophonic by nature means that spatial aspects of the original sound event will not be maintained. That is not to say that the position of the sound source would be incorrect, but rather that its physical dimensions, and the fact that the sound event occurs over more than just a single point in space, would not be accurately represented in a typical implementation of this method.

Instead of using recorded stimuli, one can also attempt to synthesize stimuli. Synthesis has the advantage that some types of variables can be controlled much more accurately than with real stimuli. Such control may be necessary to clearly establish the acoustical characteristics of a signal that are important to perception (McAdams et al., 2004), but as put by Carello et al. (2005), "The process of synthesis is a tricky one. . . " One can synthesize signals according to theoretical acoustics – idealized versions of the event that likely cannot approach the complexity of the real event – or according to perceptually motivated methods in which the stimuli carry more ecologically relevant information. Sound effects created by Foley artists are more along the lines of an information-based approach to synthesis (Heller, 2002b; Carello et al., 2005). Both methods may have their place, but accurate synthesis apparently requires consideration, from the beginning, for the complexity of real physical events. It can be seen from the research described in this chapter, that constrained synthetic events are often the ones most difficult for listeners to extract information from (e.g., Halpern et al., 1986). Those stimuli that the researchers may themselves have difficulty describing, mathematically or theoretically, have often been the easiest for listeners to extract information from.

For any of these methods, the headphones or earphones used for playback should either have a flat frequency response, or an equalization should be done before the signals reach the headphones so that the influence of the headphones does not distort the information. Few of the research experiments described in this chapter have included consideration for this fact, a notable exception being that of Ishibashi & Preis (2005a,b). Researchers have perhaps considered the frequency response of their headphones to be sufficient or have been unable to perform an equalization due to a lack of resources.

While issues such as the equalization of the frequency response of headphones *may* be assumed to be of little influence, one should be confident of this or have other good reason to believe the procedure is not necessary before skipping it or any other steps. It seems reasonable that if one is developing theories or models of how a particular sound-producing source attribute may be perceived, then one should first show that any assumptions being made are reasonable ones. For example, if one presents monophonic recordings to test subjects instead of presenting the stimuli live, it would be wise to first demonstrate that the information lost by doing so is not important to perception.

When selecting headphones for a listening test, experimenters must choose whether they will use so-called “open” or “closed” headphones. In addition to possible technical performance differences, closed headphones attenuate sounds from the external environment. This may make the wearer uncomfortable because of the inability to perceive his or her auditory environment, much like when wearing earmuff-type hearing protectors. Open headphones attempt to minimize this, but of course are thereby unable to attenuate *unwanted* noise, if such noise is present in the test environment.

Classically, many psychoacoustic listening tests have been conducted in either anechoic chambers or other special booths designed to attenuate the sound from outside of the booth. These booths are also typically designed to be acoustically dry – that is, an attempt is made to attenuate reflections from walls. On first thought, one might assume that reverberation may make a listening task more difficult for test subjects, but this may not necessarily always be true. For example, early reflections have been shown to aid in speech intelligibility tests by increasing the effective signal-to-noise ratio of the signal (Bradley, Sato & Picard, 2003). Listening booths are very useful for controlling listening conditions. However, their generally small size is also misrepresentative of most rooms that listeners would be in during their daily lives. Testing in larger, acoustically normal rooms further enhances the ecological validity of listening tests by providing test subjects with a comfortable environment and one in which they are used to hearing sounds. Although they are not as frequently built as listening booths, larger rooms may also be designed to attenuate noise coming from outside of the room. Such rooms may be an excellent choice for those conducting everyday sound research, as they could provide both a natural environment and control of external noise, simultaneously.

It is also possible to present stimuli live during listening tests. Live presentation has nearly been forgotten after the advent of signal generation and recording techniques, but in many ways, it is superior to any of the methods described above. Listening tests in which stimuli are presented live *may* be more difficult to set up and perform than tests based on pre-recorded stimuli, but they are likely to be far more perceptually realistic. The additional time and cost required to produce apparatuses for presenting live stimuli has been suggested as the reason that more live studies are not conducted (Lutfi et al., 2005). An additional potential disadvantage is the fact that not all test subjects may hear strictly identical versions of the stimuli. However, whether or not this is actually a disadvantage depends on the research questions. As an example, consider an experimenter who decides to use recordings for listening tests. In preparation for the tests, many stimuli are recorded, and the experimenter selects just a few of them for use in the test. In this case, the experimenter is relying on the fact that he or she has not mistakenly chosen recordings of “outlier events” or recordings of somehow physically and acoustically awkward nature that may have occurred due to random variations that may exist between the events. If live presentation is used, there may also be outlier events, but because the same signals are not heard by all subjects, the presence of one will not end up influencing all subjects as it would if such an erroneous event had been selected from the recordings for presentation to all subjects. This hypothetical situation should serve to point out that the fact that if each listener does not hear precisely the same stimuli as every other listener, it may not be a complete disaster. It could be a benefit.

If one is interested in determining the capabilities of human listeners to gather information from sound events, live presentation may be the optimal choice. If one is interested in forming an acoustic description of the perceptually salient features of everyday sounds, some form of recording may be necessary, either simultaneously at the time of live presentation during a listening test or for the purpose of conducting the listening test from the recordings. The first method, in which live presentation occurs while an acoustic manikin (for example) also records the stimuli from a position near the subject, would presumably be an effective way at both presenting ecologically valid stimuli to the listener as well as simultaneously documenting the signals for later analysis.

Some may complain that when using live presentation, researchers are limited in their abilities to fully control all parameters being presented to subjects. Such complaints may be guided by hypotheses that certain acoustical variables are responsible for the perception of, for example, size. It may be true that a single acoustical variable may not be adjustable while leaving all others constant, but if such parameters cannot be created with live stimuli, consideration should be made of whether or not they are even worth testing. While live presentation may not allow fine control of acoustical variables, live presentation allows great control over perceptually-relevant source properties. For example, the material of an object can be changed while keeping its size fixed, and vice versa. Live presentation restricts the type of stimuli to those that are physically possible. There is never any doubt that the stimuli used in tests involving live presentation are accurate representations of real events.

2.3.2 Response Methods

Many researchers have demonstrated that active response tasks provide better estimates of visual distance perception than do traditional magnitude estimates (e.g., Loomis, Fujita, Da Silva & Fukusima, 1992; Rieser, Ashmead, Talor & Youngquist, 1990; Bootsma, 1989). Witt, Proffitt & Epstein (2005) found that reachability judgments were influenced when the test subject intended to actually reach as a part of the task, but not so when the actual reaching was not part of the task. Similarly for auditory tasks, others have shown that action-based responses, in which a subject was required to walk to the source of a sound, were more accurate than estimates of feet, inches, angles, or other non-action based methods (Ashmead et al., 1995; Russell & Schneider, 2006; Neuhoff, 2001a). There is therefore formal evidence that, at least for perception, "actions speak louder than words."

There is also evidence that perceptual judgments become more accurate when subjects are prevented from becoming excessively analytical in their responses. Heft (1993) performed investigations in which it was shown that when subjects were asked to estimate distance in a reachability experiment. Subjects did this as a primary task in one experiment and as a subsidiary task in another experiment. Presumably, subjects were less analytical about their judgments of reachability (i.e., distance) when they were concentrating on performing a different primary task (solving a puzzle) than when the judgment was performed as the primary task. The results indicated that subjects were significantly more accurate in their judgments for the condition in which the reaching judgment was made as a subsidiary task.

A third condition in which the subjects made reachability judgments as a primary task, but in which their time was limited, was also completed. The results appeared to indicate that subjects were better than when there was no time limit, but the results were not statistically significant. However, the results were also not found to be significantly worse than the condition in which the subjects made reachability estimates as a secondary task. The main lesson from this experiment was that perceptual judgments made by test subjects given excessive opportunity to contemplate their judgments may be worse than more natural judgments as would be made in the real world.

Though it has been shown long ago that humans are able to echolocate surfaces (e.g., Supa, Cotzin & Dallenbach, 1944), it was unclear until recently if self-motion would improve performance. Ashmead et al. (1995) showed that auditory distance perception is more accurate when the listener is walking while listening to the stimulus as opposed to standing still while listening to the stimulus. More recently for echolocation tasks, it has been shown by Rosenblum, Gordon & Jarquin (2000) that subjects could attain slightly better accuracy if they made their judgments while moving. Considerations were made to account for the fact that this was probably not simply due to the listener being able to echolocate from more than one fixed position, but that the relationship between position and motion was important.

For a distance estimation task, Witt, Proffitt & Epstein (2004) found that as the effort involved to complete the task increased, perceived distance also increased. When subjects were asked to throw a ball three times to a particular distance, and then verbally report the distance to which they were attempting to throw, perceived distance was greater for heavy balls than for lightweight balls. There was no difference between the distances to which the subjects actually threw the balls. Similarly, if listeners were first asked to throw a heavy ball three times, verbally report distance, and then either walk to the target or again throw the heavy ball to the target, subjects' estimated distances were apparently influenced by their prior experience with throwing the ball. Their estimates were significantly longer when they intended to throw the ball than when they intended to walk to the target. Apparently, the energy required to complete a task influences the outcome of the task (Proffitt, Stefanucci, Banton & Epstein, 2003). Thus, the role and involvement of the test subject appears to play an important part in the results. Perception is apparently influenced by both the environment and the subject making the perceptual judgments. The actual physical parameters, the task, and the effort required to perform the task all play a role in the reported perception (Witt et al., 2004). Therefore, it appears to be important to consider experimental response methods carefully.

2.4 Conclusion

Gibson (1966) argued that perceptual research in general has focused excessively on a micro-level, the variables of which may not be relevant for perception as it relates to behavior. Plomp (2002) has argued the same, specifically for research on audition. Excessive emphasis has been placed on explaining perception using the variables of physics, with no constraints on which of those variables may be relevant to real-world perception (Schmuckler, 2004). The psychoacoustician's library of physical variables and the sensations produced by them have proven inadequate and incapable of explaining perception. A neat correspondence between the receptors of perceptual systems and the variables of physics may not exist (Gibson, 1966), and there is evidence that quantities such as just-noticeable-differences for independent sensations cannot be used to predict the perception of source characteristics such as material (e.g., Lutfi & Oh, 1997). According to some (e.g., Carello et al., 2005), it is entirely possible that existing ways of describing information may not be sufficient for describing parameters important to perception. Intuitively relevant perceptual properties such as heaviness or loudness as estimates of physical mass or sound pressure level may not even be appropriate for describing *behaviorally* relevant characteristics of objects in the world. Supposed "errors" in perceptual judgments may therefore not reflect an error on the part of the subjects, but an error on the part of the experimenter who has not asked questions of significance to a person who is used to perceiving and acting in the world (Warren, Jr. et al., 1987).

It is clear that static test environments using artificially simple stimuli and focusing on studying sensations cannot be used alone as a way to understand a perceptual system. This is a point that is broadly accepted by many researchers in the field of visual perception (Neuhoff, 2004c). Auditory research must embrace the fact that the world in which its test subjects live is rich with information, and it is the purpose of the auditory system to take advantage of this information. Much of the research on everyday listening that has been described in this chapter can be used as a guide for exploring the mechanisms responsible for information perception more deeply.

Chapter 3

Hearing Impairment and Hearing Aids

> *If it were easy to detect pure sensations, we could all be representational painters without training.*
>
> -James J. Gibson (1966)

3.1 Hearing Impairment

Figure 3.1 shows a diagram of the anatomy of the peripheral ear. For discussion purposes, the parts of the ear shown in the diagram are commonly divided into three main parts. The pinna and ear canal are referred to as the *outer ear*. The tympanic membrane (eardrum) serves as the division between the outer ear and *middle ear*. The *middle ear* consists of the tympanic membrane and three bones called the malleus (hammer), incus (anvil), and stapes (stirrup). The cochlea is located in the *inner ear*, and produces electrical activity, which is passed by the auditory nerve to the processing centers of the brain.

Sound waves that pass through the ear canal reach the tympanic membrane, which is the first part of a stage that converts pressure variations in the air (the acoustic potential) into fluid vibrations in the cochlea. The bones of the middle ear act to efficiently transfer the vibrations in the air to the fluid-filled cochlea, where the vibrations are converted to nerve spikes.

The cochlea consists of various membranes that vibrate as "pushed" to do so by

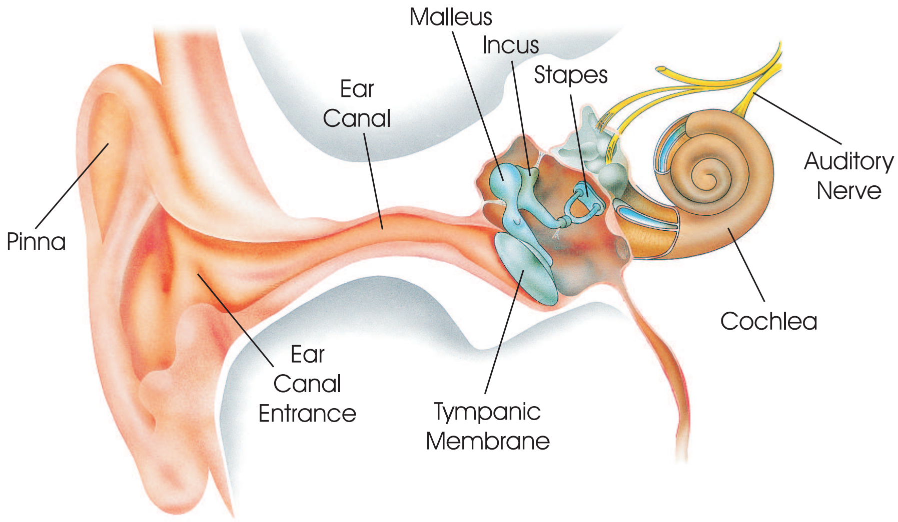


Figure 3.1: Anatomy of the peripheral ear. Illustration provided by Oticon A/S.

the vibrations reaching the cochlea via the bones of the middle ear. On one of the membranes inside the cochlea, the basilar membrane, sit rows of hair cells. These hair cells are critical to the transduction of vibrations into nerve impulses. There are two types of hair cells, *inner hair cells* and *outer hair cells*. The role of the inner hair cells is to help perform the actual transduction of vibration to nerve impulses to the afferent pathway of the auditory nerve (i.e., toward the brain). The outer hair cells help to mechanically amplify the vibrations produced by quiet sounds so that the inner hair cells have a large enough input to detect. Their function is nonlinear in that they provide significant amplification to very quiet sounds, but have essentially no effect at higher sound levels (above approximately 60 dB).

The most common measurement used to diagnose hearing problems is the pure-tone air-conduction audiogram. A variety of measurements may typically be made on a person seeking help with a hearing problem, but the audiogram is that which is most often used to describe the hearing loss. A pure-tone air-conduction audiogram is a measure of the quietest pure tone a listener can hear for a series of frequencies. The audiogram is typically measured for nine frequencies or fewer between 125 Hz and 8 kHz. The most-comfortable levels and uncomfortable levels are also often measured to give an audiologist or hearing specialist a better understanding of the hearing impairment. Speech intelligibility tests may be conducted as a measure more realistic of real-world hearing ability. Other techniques such as bone-conduction threshold measurements, recordings of otoacoustic emissions, and auditory brainstem response measurements may also be used to diagnose the nature of hearing problems.

The most commonly diagnosed hearing losses are categorized as *conductive* losses and *sensorineural* losses. Conductive losses are typically caused by fluid in the middle ear, a perforated eardrum, a growth in the ear canal, excessive earwax blocking the ear canal, or by a reduction in the mobility of the middle ear bones (otosclerosis). Otosclerosis is caused by an excessive build-up of bony material that prevents the middle ear bones from moving as freely as they normally would. These types of hearing losses are generally fairly easy to treat.

Sensorineural hearing losses are the most common type of hearing loss, and unfortunately, they are also often difficult to treat. They can originate in the cochlea or auditory nerve, and may be a result of a variety of causes including acoustic trauma, infections, the use of ototoxic drugs, inner-ear diseases like Morbus Ménière, or more commonly presbyacusis (age-related changes) or genetic predisposition (Pickles, 1988). *Retro-cochlear* hearing impairments are caused by damage to the auditory nerve or otherwise beyond the cochlea, but the majority of sensorineural losses are believed to be the result of damage to the outer hair cells in the cochlea. Damage to the outer hair cells results in an inability for sufferers to hear quiet sounds, as the mechanical amplification mechanism is damaged. The range of sound levels audible to a person suffering from a sensorineural hearing loss is therefore reduced compared to the range of sound levels audible to a normal-hearing person. At levels just above the raised threshold of the hearing-impaired listener, increments in sound level result in abnormal increments in perceived loudness. Loudness increments are perceived as being larger just above the raised threshold, but usually become normal at high levels. This phenomenon is known as *loudness recruitment*. Many people suffering from sensorineural hearing impairment experience this reduction in the dynamic range between the just-audible threshold and the threshold of discomfort. Damage to inner hair cells can additionally prevent the sufferer from hearing even moderately loud and loud sounds.

An audiogram by itself provides only a rough description of the ability of the subject to hear in the real world. Sufferers of sensorineural hearing loss may have many other problems in addition to a simple reduction in the dynamic range of sound levels that they can perceive. Classical measures frequently indicate a decreased ability to distinguish between frequencies, a worsened ability to separate sounds in time, and more trouble in perceptually separating sounds that occur simultaneously. These problems are currently more difficult to treat than “simple” audibility issues. A fundamental part of the hearing loss, the decreased dynamic range of acceptable sound level inputs, is the part that is most frequently dealt with by hearing aids attempting to provide help to those suffering from hearing losses. Hearing aids and cochlear implants are the two most common techniques for helping the hearing impaired, but cochlear implants are typically only used to aid those with severe hearing losses. Tactual aids in which acoustic information is translated into patterns of vibration on the skin of the user are also an option for severely impaired subjects, but hearing aids are far more common and will be further discussed.

3.2 Hearing Aids

In simple terms, a hearing aid is a device to amplify the sounds reaching a listener to a level that is appropriate for the listener. The hearing aid makes quiet sounds audible and attempts to present the full range of sounds from the world in the compressed dynamic range of the hearing-impaired listener. Because of the nonlinear nature of common, sensorineural hearing losses, an amplifier that simply makes all sounds louder, regardless of their original level, is inappropriate. While a person suffering from a sensorineural hearing loss may require soft sounds to be amplified by a great deal, they often do not require as much amplification for loud sounds. An amplification system in which consideration is made for this is therefore required.

A *compressor* is used in conjunction with an amplifier to vary the gain depending on the input level. A simple hearing aid compression system in which the gain is 30 dB for input levels lower than 50 dB, and above which the gain slowly decreases, is shown in Figure 3.2. The plot shows the amount of gain provided by the hearing aid for a range of possible input levels. The amount of gain is chosen in an attempt to make previously inaudible sounds audible and of a comfortable level. Many fitting rationales exist with different ideas of the most appropriate way to prescribe gain. For example, some seek to maximize speech intelligibility while others seek to restore the sensation of loudness to the level that would exist for a normal-hearing person. It seems that none of the methods are completely capable of predicting the optimal gain for all users, a fact which may indicate that either there are insufficient ways of characterizing hearing impairment or that the data available is not being used effectively.

A compression system can also be described by an input/output curve. An example input/output curve for the same compression system depicted in Figure 3.2 is shown in Figure 3.3. For this particular figure, the amplification is linear for input levels less than 50 dB, that is to say, for an increment in input level of 10 dB, the output level is also incremented by 10 dB. The slope of the input/output curve below 50 dB is said to be 1:1. Above 50 dB, the compressor is active and no

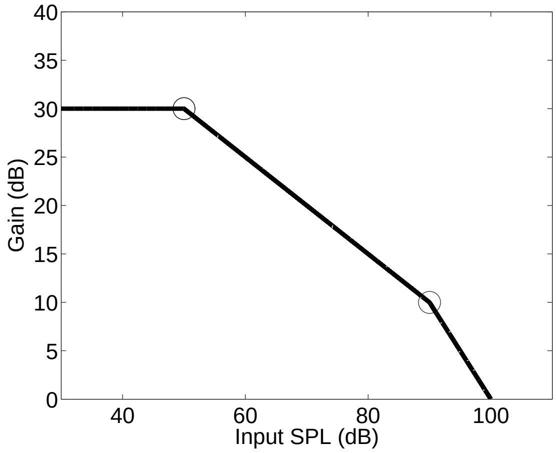


Figure 3.2: Gain (dB) versus input sound pressure level (dB) for a hypothetical nonlinear hearing aid in which 30 dB of gain has been applied for quiet sounds. The amount of gain gradually decreases for louder sounds.

longer amplifies by the same amount. The point at which this change occurs is commonly called the compression threshold or the knee point of the compressor. For this example, sounds above 50 dB are compressed at a ratio of 2:1. That is, for an input increment of two units, the output is only raised by one unit. In the compression system shown in Figure 3.3, a maximum power output has been set at an output sound pressure level (Output SPL) of 100 dB.

Because compressors cannot act instantaneously, the input/output curve describing its operation is technically only valid for steady-state conditions. The compressor generally requires, and is programmed to have, a reaction time. If the hearing aid is exposed to a stationary, quiet background noise and suddenly the level rises, the compression system takes time to react. This time is referred to as the *attack time* of the compressor. With a sudden increase in input level, the gain will initially remain the same as before the increase, but it will gradually decrease until it reaches the steady state as defined by the input/output function. In rough terms, the attack time specifies the time it takes for the compressor to reduce the gain so that the output level specified by the input/output function is reached.

If a hearing aid is exposed to a constantly loud sound and then suddenly to a softer sound below the compression threshold, the hearing aid will also take time to reach its steady-state prescribed gain. Because the hearing aid has been in a mode in which it is prescribing relatively little gain, a sudden decrease in input level will result in less gain provided to the soft sounds than would normally be provided in the steady-state condition. The time required for the hearing aid to

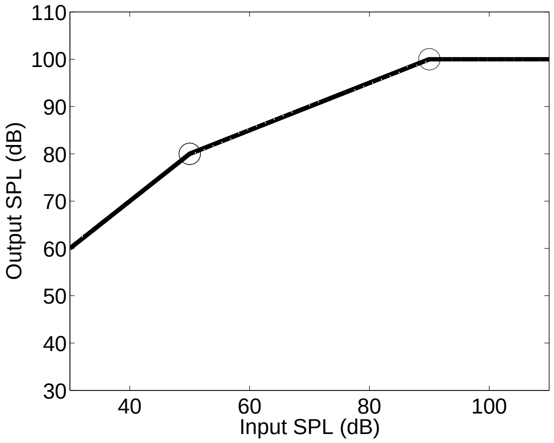


Figure 3.3: Sample input/output curve of a nonlinear hearing aid. Output sound pressure level is related to the input sound pressure level in a nonlinear fashion. Absolute output level for a particular input level depends on the prescribed gain. Compressor knee point at an input level of 50 dB is circled. Maximum power output at an output level of 100 dB is also circled.

reach the steady-state condition is called the *release time*.

There is no consensus on the optimal settings for attack and release times. Either may be limited for technological reasons (i.e., a compressor cannot act instantaneously) or by design. Optimal settings likely depend on the sound environment of the wearer, cognitive skills of the wearer (Gatehouse, Naylor & Elberling, 2003; Lunner, 2003), and other factors. Generally, the settings are optimized according to the ways that are believed to provide the wearer with good speech intelligibility and comfort. Suggested release times may vary between fitting rationales anywhere from tens of milliseconds to a few seconds. The effect of compression on the input signal is greater with high compression ratios and short time constants.

In simple implementations, a single compressor may be used to compress the dynamic input range for all input frequencies, but the input signal may also be split up into multiple channels. In such a case, each channel would cover a range of frequencies, with the compressors of each channel able to have different compression ratios, attack times, and release times. Similarly, more complex hearing aid compression systems compress different input levels by different amounts. Whereas the hypothetical compression system shown in Figures 3.2 and 3.3 has only two compression regions below the maximum power output level, modern hearing aids may have more. Each of these compression regions may have their own compression ratio. Some of the hearing aids used in the experiment described in Chapter 5 have four compression regions below the maximum power output level.

Hearing aids will also have a maximum amount of power that they can deliver. An output limiter is often used to control the sound at levels near the maximum output level. As an example, if the maximum output level of the hypothetical hearing aid depicted in Figures 3.2 and 3.3 was 100 dB, input sound levels of greater than approximately 90 dB would not be provided with any amplification. The output limiter is essentially a compression system with a very high compression ratio and a fast attack time. Note that the output limiter may not work quite as abruptly as illustrated here, but will likely gradually limit the output over a range of a few decibels in input level. Such an output limiter may be in place in order to prevent sound levels from becoming uncomfortably loud or because of practical limitations in the hearing aid. The limiter may be necessary because the hearing aid is unable to produce sounds louder than 100 dB (a hypothetical value for this example, but which may vary for real hearing aids) and to avoid unpleasant distortion as the output reaches this level. However, by its very nature, the output limiter will result in another type of distortion: the distortion of the envelope of the signal being processed by the hearing aid.

Hearing aids come in a variety of physical shapes. The most common types are:

- Behind-the-ear (BTE),
- In-the-ear (ITE),
- In-the-canal (ITC),
- Completely-in-the-canal (CIC), and
- Open fittings.

Essentially, these categories describe the way in which the hearing aid is worn. Behind-the-ear aids are worn, as their name implies, behind the ear and are typically able to be larger than hearing aids inserted into the ear or ear canal. Their size allows them to accommodate electronics for large amounts of amplification, and for this reason, they are often used when large amounts of gain are needed. Completely-in-the-canal aids are typically the smallest and have cosmetic advantages for this reason. With the exception of open fitting hearing aids, the above types of hearing aids typically require an impression of the wearer's ear to be made. For in-the-ear aids, in-the-canal aids, and completely-in-the-canal aids, the hearing aid is built into a custom shell that is made from an impression of the wearer's ear. A behind-the-ear aid is usually attached to an individually shaped earmold via a tube. Behind-the-ear hearing aids were used extensively in the experiments described in Chapter 5 and will therefore be described in greater detail here.

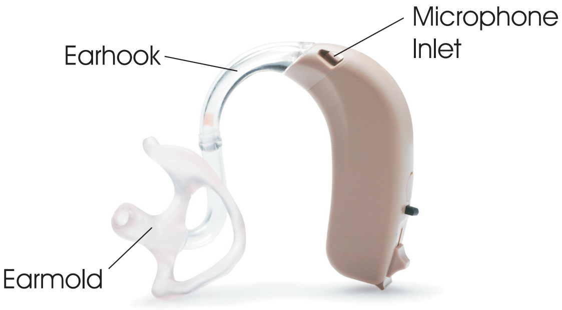


Figure 3.4: Behind-the-ear hearing aid attached to an earmold. Photo provided by Oticon A/S.

Figure 3.4 shows a picture of a behind-the-ear hearing aid. The microphone of a behind-the-ear hearing aid typically sits just above the pinna when the hearing aid is worn. The ear hook leads from the hearing aid towards the entrance to the ear canal, where it connects to an earmold inserted into the wearer's ear canal. A behind-the-ear hearing aid has a disadvantage, when compared to some other types of hearing aids, in that the microphone is placed above and outside of the wearer's pinna. Because of the microphone placement, behind-the-ear aids effectively distort the spatial sound field that would otherwise be better preserved if the sound were picked up within the confines of the pinna. In-the-ear, in-the-canal, and complete-in-the-canal hearing aids are equipped with microphones that are located much closer to the entrance to the ear canal, and thereby they are better able to take advantage of the directional cues provided by the pinna.

Modern hearing aids of all styles may typically have a number of other features that will not be described in detail here. These include noise suppression systems, directional microphones that may attenuate sounds coming from directions other than the front of the wearer, and feedback cancellation systems that enable the hearing aid to produce a higher amount of gain without resulting in acoustic feedback. All of these features risk introducing distortion or even do so purposefully with the goal that it will help the wearer.

In general, modern hearing aids often pack a great amount of processing into a very small space, but as would be expected, the audio quality may not be that of a high fidelity home stereo of equivalent cost. Hearing aid designers are frequently faced with construction issues on which they must make compromises in audio quality. As an example, typical hearing aids do not provide much amplification beyond 8 kHz. A reduced frequency bandwidth or a higher amount of distortion may be accepted by the designer in order to save space. The impact of such distortion and bandwidth limits on the perception of everyday sounds is largely unknown.

3.3 Hearing Impairment and Information-Based Perception

It could be argued that hearing threshold measurements such as audiograms say very little about hearing losses, but they are often used as the basic data for fitting hearing aids. They are even only partially effective in predicting speech intelligibility performance (Jenkins, 1985). With the exception of speech perception, little is known about the abilities of the hearing-impaired population to gather information from other types of sounds. It could very well be the case even that the processing in hearing aids, which is designed to improve speech intelligibility, hinders the ability of the wearer to obtain information from everyday sounds.

Psychoacoustics has responded to some complaints made by normal-hearing and hearing-impaired listeners. The responses primarily address complaints concerning noise, comfort, and problems for the hearing impaired to understand speech. Perhaps because the hearing-impaired population may not typically complain about an inability to hear everyday sounds or perhaps because interest in acoustic signals other than speech and synthetic sounds is new, little research has been done on how well everyday sounds are perceived by hearing-impaired listeners. However, it could just be that no one has asked the hearing impaired about their abilities to hear everyday sounds.

It has been suggested that a reduced ability to perceive environmental sounds can lead to a feeling of detachment and uncomfortableness (e.g., Ramsdell, 1978). In a brief study, Pichora-Fuller (1999) concluded that there were many aspects of one woman's sound environment that were important to her, but that hearing rehabilitation techniques were not doing a good job at addressing these issues. It is likely that the single test subject, who provided a subjective assessment of the importance of sounds in her daily life, is not unique. Noble (1983) stated that, at least at the time of his writing, hearing aids fail to account for the real world. As evidence of this, the author cited the fact that many hearing aids go unused or are used only in limited amounts. While many advances have been made since his words were published, they may serve as a reminder that hearing aids must account for, and be designed for, the real world if they are to be of optimal benefit.

Hearing aid fitting typically revolves around an attempt to improve speech intelligibility for the wearer of the hearing aid. With the occasional exception, little thought is normally given to the multitude of other sounds important for a person to operate and be comfortable in his or her daily life. Gatehouse (2005) examined correlations between a measurement of degree of handicap and the subscales of a questionnaire that contributed to the measure. The Speech-Hearing, Spatial-Hearing and Qualities of Hearing Questionnaire (SSQ) was designed in an attempt to investigate aspects contributing to hearing impairment as a handicap by examining issues which are typically not the focus of research and are not typically reported by those suffering from hearing impairment. He found that topics, which have historically been of great focus in hearing-impairment research, such as speech in quiet and speech in adverse conditions, were not the most important factors contributing to the feeling of hearing impairment as a handicap. Topics such as the identification of sounds, perception of distance and movement, listening effort, speech in speech maskers, and multiple stream processing were found to be of greater influence. Apparently, informational aspects of sound events other than speech are also worth addressing.

The pressure variations resulting from an event and thereafter propagating through the air in the environment of a listener can be considered as a potential stimulus. Whether the stimulus can be called a "sound" or not depends on the receiver (Gibson, 1966). A potential stimulus may provide one set of information to one person, but a different set of information to another person. An information-based approach to hearing-impairment research should consider, as a founding point, that the hearing impaired are not as capable of extracting information from potential stimuli as are the normal-hearing population. For people who have lost their hearing, the world is altered when the information that was formerly available to them is no longer available. An inability to understand speech is apparently a major complaint in this situation, but as described above, it is possible that everyday sounds are also likely an important part of human lives.

Although some acoustic stimuli may still be audible, the information may not be as easy to extract for a hearing-impaired person as it is for a normal-hearing person. It is well known that the hearing impaired are often not as capable of dealing with degradations to the quality of speech as are normal-hearing people (e.g., Plomp, 1977; Gelfand, Ross & Miller, 1988; Soli & Nilsson, 1994; Cox, Gray & Alexander, 2001; Schneider, Daneman & Pichora-Fuller, 2002). While normal-hearing people are amazingly capable of using redundant cues in speech and extracting invariants even in acoustically complex situations, hearing-impaired listeners often suffer – both with and without hearing aids. While this is likely to be the case also for everyday sounds, this fact has not yet been properly addressed by science.

Few studies have examined the abilities of hearing-impaired listeners to gather information from everyday sounds. A few investigators have asked cochlear implant recipients to identify environmental sounds (e.g., Tye-Murray, Tyler, Woodworth & Gantz, 1992; Proops, Donaldson, Cooper, Thomas, Burrell, Stoddart, Moore & Cheshire, 1999; Reed & Delhorne, 2005). They all found a wide range of abilities. Some subjects performed very well, while others performed very poorly. Reed

& Delhorne (2005) suggested that temporal cues were largely responsible for test subjects' abilities to identify the sounds in their study. This is perhaps to be expected since cochlear implants do not transmit much spectral detail. It was hypothesized that subjects who performed well may have taken advantage of gross spectral cues preserved by the cochlear implants.

In an information-based approach to correcting hearing impairment, a focus is placed on improving the coupling of information from a person suffering from a hearing impairment and the environment in which the person lives and acts. Rehabilitation may be helped in many ways, not only by prostheses for the hearing impaired but also by changes to the environments and players in the environments of the hearing impaired (e.g., Noble & Hétu, 1994; Borg, 1998). In an ecological context, Noble & Hétu (1994) suggested that a hearing aid or other prosthesis may be considered as one way (of many) to improve the flow of information between the user and his or her environment.

If it is ultimately desired to reduce to a minimum the influence of hearing impairment as a handicap, ways must be found to enable the hearing impaired to acquire the same information from their environments as do the normal hearing. A rough process to accomplish this goal, using assistive listening devices, is proposed here:

1. Determine if there is a difference between the amount or accuracy of information that a hearing-impaired person gets from his or her environment compared to a normal-hearing person.

2. If the hearing impaired are not capable of obtaining the same information from their environments as the normal hearing can, identify which environmentally relevant information is missing or altered for the hearing-impaired population. Evaluate the practical significance of these differences with a consideration for the sizes of the differences.

3. Focusing primarily on the information that is missing for the hearing impaired, determine by which means the information is normally provided to the normal hearing.

4. With knowledge from the previous steps as a guide, identify the ways in which the hearing impaired are different from the normal hearing.

5. Determine which of the hearing-related differences between the hearing impaired and the normal hearing contribute to the decrease in accuracy or the decrease in quantity of perceived information.

6. Consider how information transmission may be restored for the hearing im-

paired.

  (a) Determine if current solutions, such as hearing aid sound processing techniques, are disrupting information transmission.
    i. Examine whether the assistive listening device or the way in which it is used is responsible for eliminating important parts of the incoming information prior to the device even beginning its processing (e.g., spatial information may be lost by fitting a hearing aid on only a single ear of a person with a bilateral hearing loss or simply due to the placement of the hearing aid and its microphone or microphones).
    ii. Check whether intentional distortion (e.g., compression) or unintentional distortion (e.g., harmonic distortion) introduced by the device is compromising information transmission to the wearer.
  (b) Assess what is currently being done to aid the hearing impaired in receiving information.
  (c) Determine what can be done to restore the coupling of information between the hearing impaired and their environments to a level like that of the normal-hearing population.

Each of these steps likely requires a great deal of work. For example, it may be a large undertaking to determine which differences between the hearing impaired and the normal hearing contribute to the worsened performance of the hearing-impaired (step 5). In speech research, this issue remains unsolved, with apparently many factors influencing speech perception. It should also be kept in mind that a prosthetic device may not be capable of completely restoring hearing to normal levels, and therefore simultaneous efforts are required on other fronts. Also, note that according to the steps outlined in this plan, it is not immediately obvious that there is ever a need to understand musical listening (i.e., sensations) for either the normal hearing or the hearing impaired. An understanding of how ecologically relevant information is transmitted and perceived by listeners is important.

Whether it is even possible to enable the hearing impaired to obtain the same information from their environments as do the normal hearing is an issue that will not be discussed here. There very well could be technological restrictions or conflicting interests that prevent complete restoration of everyday sound information, but for the purposes of this project, all feats will be assumed attainable. Discussing potential roadblocks at this point would surely serve to limit potential solutions.

Chapter 4

The Influence of Stimulus Presentation Method on Auditory Perception of Object Length

> *I can imagine a conference of "bat psychologists" asking how it is that human beings transform visual information into sound so they can navigate in the truly auditory world!*
>
> -James J. Jenkins (1985)

This chapter is based on, and builds upon, articles written for the Twenty-First Danavox Symposium and the Thirteenth International Conference on Perception and Action (Kirkwood, 2005a,c). Posters, summarizing the articles, were also presented at each meeting (Kirkwood, 2005b,d). Additionally, the experiments were described in an article written for the course *Writing & Reviewing Scientific Papers* held at Aalborg University in Aalborg, Denmark (Kirkwood, 2005e).

4.1 Introduction

Humans have been shown to be able to hear the relative sizes of speaking humans (Smith, Patterson, Turner, Kawahara & Irino, 2005). Evidence for size information in vocal sounds has also been demonstrated for other animals such as monkeys (e.g., Fitch, 1997), and it has been suggested that animals use information concerning the size of other animals to determine with which animals they will communicate (e.g., Narins & Smith, 1986). It appears from these and other studies that there is good reason to suspect that size perception, and auditory recognition of size in particular, is an important part of mating, survival, and the behavior of an organism in its environment.

In addition to size information in vocal sounds, there is evidence of size information in everyday sounds. Size, shape, and position have been shown in numerous studies to be perceivable from everyday sounds by normal-hearing listeners, solely on the basis of sound (e.g., Carello et al., 2005; Grassi, 2005; Heine, Guski & Pittenger, 1993b). In particular, the sound resulting from an object impacting another object has been shown to be informative about size and many other properties of the objects involved in an impact. Normal-hearing people have been shown to be capable of hearing the sizes of dowels dropped on a floor (Carello et al., 1998), about the relative sizes and shapes of struck plates (Kunkler-Peck & Turvey, 2000), about the sizes of balls dropped on plates (Grassi, 2005), about the materials with which struck objects are made (Giordano & McAdams, 2006), and other properties.

In classical psychoacoustic experiments, monophonic signals are used for presenting stimuli to test subjects. In the more ecologically-motivated listening tests as discussed in this thesis, experimenters have used a variety of stimulus presentation methods, including headphone presentation of monophonically recorded stimuli (e.g., Warren, Jr. & Verbrugge, 1984; Halpern et al., 1986; Fowler, 1990; Rosenblum et al., 1993; Houben, 2002; Grassi, 2002; Tucker, 2003; Lutfi et al., 2005; Giordano, 2005; Giordano & McAdams, 2006), binaural presentation of stimuli recorded with acoustic manikins (e.g., Heine et al., 1993c; Sandvad, 1999; Ishibashi & Preis, 2005b; Robart & Rosenblum, 2005b), and live presentation (e.g., Rosenblum et al., 1996; Carello et al., 1998; Kunkler-Peck & Turvey, 2000; Grassi, 2005). Live presentation of stimuli, as has been done in some of these experiments, has the great advantage of bypassing any limitations of a recording and playback chain and can thereby increase the ecological validity of a listening task. Recorded stimuli may not be as perceptually realistic to test subjects, and therefore they may be perceptually inaccurate. However, live presentation may be difficult to implement, and this ecological validity comes at a cost in the form of a significant reduction in test method options and in the types of stimuli that can be presented. Theoretically, it should be technically feasible to produce recordings that include nearly all of the information contained in an original acoustic stimulus, but it is unclear what the influences of recording imperfections are on the perception of ecologically relevant stimulus attributes. An examination of the influence of presentation method could therefore be useful.

VanDerveer (1979) noticed that children listening to monophonically recorded stimuli in a sound identification task performed worse than children in a similar experiment listening to stimuli presented live. However, it was unclear from her results whether this difference was due to the fidelity of the stimuli, the attentiveness of the children, both, or other causes. In her experiments, the monophonically recorded stimuli were presented over a loudspeaker. Based on other observations made during informal tests with colleagues, VanDerveer (1979) reported that recordings made with poor equipment or not recorded with care, could lead to things not sounding "right". She wrote, "Listeners are sensitive to some very small acoustic differences among meaningful sounds."

As was described in Subsection 2.3.1 of Chapter 2, monophonic recordings are incapable of accurately representing a spatial sound field. It is hypothesized that this is an important factor to consider if one wishes to understand the perceptual capabilities of humans. The binaural recording and presentation technique (e.g., Kürer, Plenge & Wilkens, 1969; Møller, 1992) provides an attractive alternative to monophonic recordings in that it can fairly accurately maintain spatial information. However, even test subjects listening to binaural recordings have been shown to be worse at making spatial judgments than when making judgments in response to live stimuli.

In the localization experiments of Møller, Hammershøi, Jensen & Sørensen (1999) and those described by Minnaar, Olesen, Christensen & Møller (2001), test subjects were found to be worse at identifying the source, and therefore the position, of a loudspeaker producing a recording of human speech when listening to binaural recordings than when listening directly to the loudspeakers producing the speech. In this case, the recordings were made with an acoustic manikin seated in the same position in which the test subjects later sat for both the live and recording playback conditions. The researchers who conducted these experiments even went to great lengths to measure and implement custom equalization filters for individual test subjects by measuring headphone transfer functions on each subject, using small microphones placed at the entrance of their ear canals. Furthermore, the researchers tested a variety of acoustic manikins and found that none of them provided results as accurate as when the test subjects performed the test live.

In an investigation of the influence of presentation method on perceived loudness, annoyance, and the unpleasantness of a set of sounds, Çelik, Waye & Møller (2005) found that there were differences in the subjective judgments of subjects when listening to monophonic recordings played back through loudspeakers and binaural recordings played back through headphones. The sounds included recordings from a restaurant environment, recordings of traffic noise, and recordings of ventilation noise. The results indicated that the degree of loudness, annoyance, and unpleasantness for a particular sound depended on which presentation method was used, but that the trends were not identical for all sounds. For example, the monophonic technique may have consistently produced higher annoyance ratings for traffic noise but lower annoyance ratings for restaurant noise.

There are therefore both objective and subjective differences that have been uncovered when comparing presentation methods (e.g., Møller et al., 1999; Çelik et al., 2005; Guastavino, Katz, Polack, Levitin & Dubois, 2005). These findings suggest that there may also exist important subjective and objective differences when comparing presentation methods for other types of listening experiments. The presentation method chosen for a particular experiment may therefore play an important role in the outcome of the experiment.

In the present study, an attempt was made to better understand the influence of the stimulus presentation method for a typical everyday listening task. The performance of normal-hearing listeners in a dropped-rod length-estimation task, like that described in Carello et al. (1998), was compared for three cases. Subjects estimated the lengths of rods dropped on a floor by listening to 1) live presentation of stimuli, 2) headphone playback of binaural stimuli as recorded from an acoustic manikin, and 3) diotic presentation of monophonically recorded stimuli as recorded from a single microphone.

From this experiment, it was thereby hoped to determine whether the stimulus presentation method is in fact important in the perception of an ecologically relevant variable (size) for an ecologically relevant sound event (an impact sound). Because it was desired to have a result that is practically relevant, common and fairly simple recording and equalization methods were chosen so that it could be assumed that these methods could be used in future listening tests. Care was not taken, for example, to measure individual headphone transfer functions for each test subject (e.g., like that of Minnaar et al., 2001). Although this would allow for a more accurate reconstruction of the sound field (Pralong & Carlile, 1996), it may likely also be impractical for many listening tests in the future. Therefore, the present test is not of the absolute limits of the presentation methods, but to determine whether or not there are differences between practical implementations of relatively easily implemented presentation methods.

Because the stimuli chosen to be tested in this experiment have a definite physical size and because sound is radiated from more than just a single point when a rod is dropped on a surface, it is clear that an accurate reproduction of this event must account for the spatial position from which sound is naturally radiated. When a rod is dropped on a surface, sound is radiated from along the entire length of the rod, from all of its surfaces – not from a specific point on the rod. This fact, in and of itself, provides information concerning size. People may potentially make use of this information. Additionally, sound may be radiated from the surface on which the rod is dropped, depending on the properties of the rod and the surface. Therefore, in addition to the primary goal of determining if there are perceptually relevant differences between presentation techniques, a hypothesis being tested is whether or not objects with physical size can accurately be represented as point sources (as having no size) or as sources with skewed spatial information. A non-spatial reproduction method, such as the diotic presentation of monophonic recordings, would represent them in this way.

If monophonic recordings are found to be sufficiently able to produce similar results as binaural recordings and live presentation, it would be suggested that spatial information is not important to the perception of size. If differences did exist between the presentation of binaural recordings and the presentation of monophonic recordings, or between live presentation and the presentation of monophonic recordings, it may be likely that spatial cues are important to the perception of size.

4.2 Methods

4.2.1 Subjects and Procedures

Eight normal-hearing test subjects, all enrolled in an introductory psychoacoustics course, participated as paid listeners. The test subjects sat facing away from an acoustically transparent but visually opaque screen. Test subjects listened to recordings presented via headphones and to the live presentation of wooden dowels, lengths 15 to 120 cm in 15 cm increments, dropped on a linoleum floor behind the screen. The subjects produced estimates of the rod lengths by positioning a moveable surface in such a way that the distance from a fixed reference surface to the moveable surface could just be reached by a rod extended from the fixed reference surface. The moveable surface was able to slide along a track, and the subjects positioned it manually. A photo of the response station can be seen in Figure 4.1. The reporting device allowed subject responses of between 0 and 2 m, and the moveable surface was positioned at the position of longest length estimate possible prior to beginning each test. Because of the wide range of lengths capable of being reported, most subjects were forced to roll around in their chair (an office chair on wheels) during the test in order to be able to report all possible lengths. A Leica DISTO "pro^4 a" laser distancemeter automatically measured the length of each estimate. A computer then acquired and stored the data. Subjects had six seconds to respond. No feedback was given.

The test participants were told they would hear rods dropped on the floor, but were given no other information about the composition or dimensions of the rods.

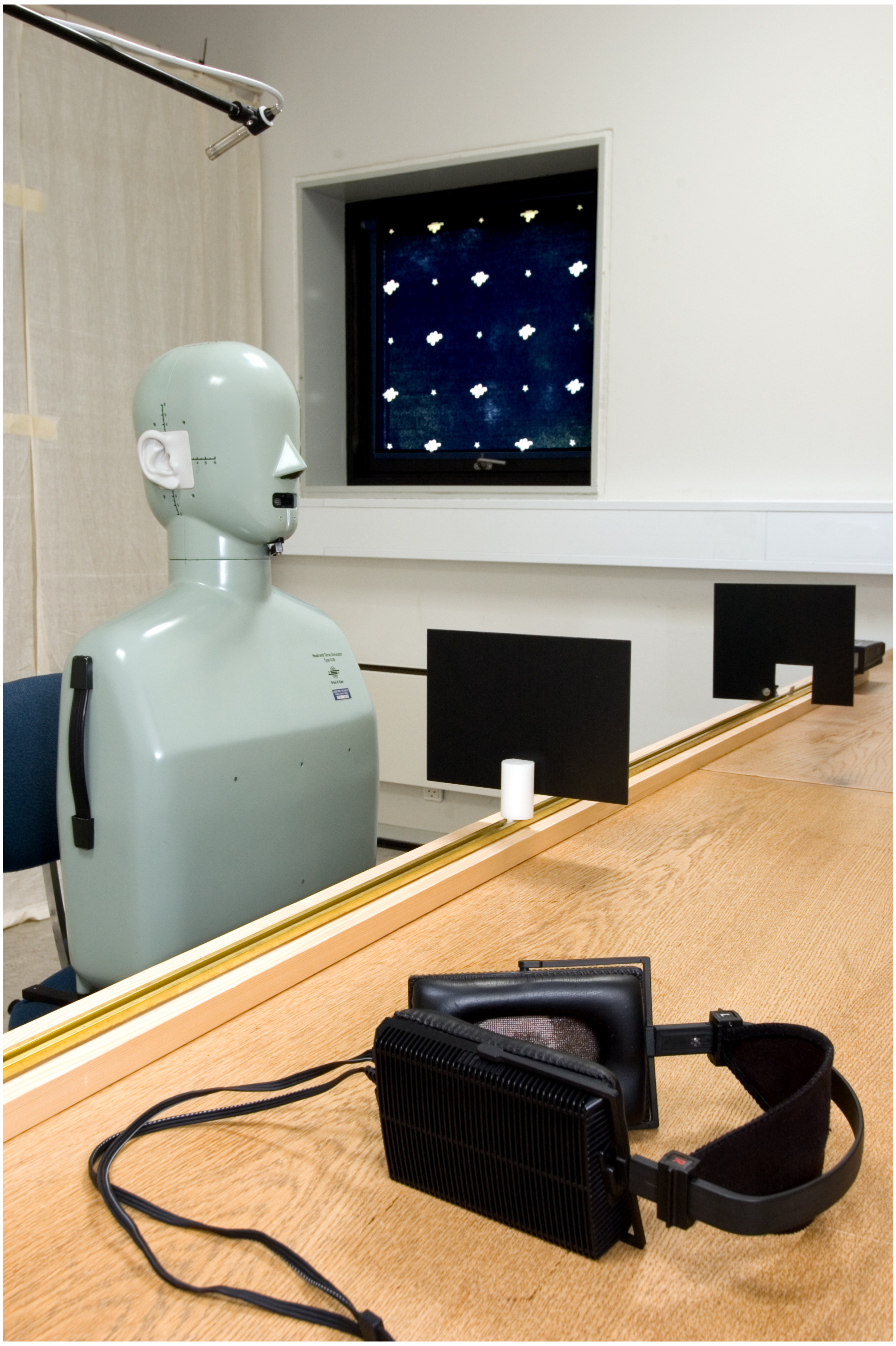


Figure 4.1: Acoustic manikin seated in the test subject's chair behind length reporting apparatus. Microphone used for monophonic recordings positioned over manikin. Headphones to be later worn by test subjects shown in foreground.

Figure 4.2: Fixture used to help drop dowels in a uniform manner on every trial. Dowel being dropped in photo is 45 cm long.

The pine dowels were dropped in a similar way on each trial: approximately 3 m behind the subject, at an angle of 15 ° relative to the floor, and with their centers of gravity positioned 59 cm above the floor. All dowels were 13 mm in diameter. In an attempt to drop them in a consistent fashion on each trial, they were released from the top of a very short (1.8 cm) ramp, which held their start position constant. Figure 4.2 is a photo of the dropping apparatus and a falling rod.

The test was broken into four sessions, one for each of the three presentation methods plus a repetition of the first session, conducted in order to enable checks of training effects. The sequence of the first three sessions was systematically randomized using Latin squares. Within each session, subjects listened to the eight different dowels being dropped onto the floor five times each. In the case of the recordings, five different recordings of each rod being dropped were used for the five replicates. Test subjects produced a unique length estimate for each drop. Eight drops, one for each rod length, were considered as a block, and all rods in a block were dropped prior to a new block beginning. To help minimize the influence of order effects, the sequence of rod drops within each block was systematically randomized.

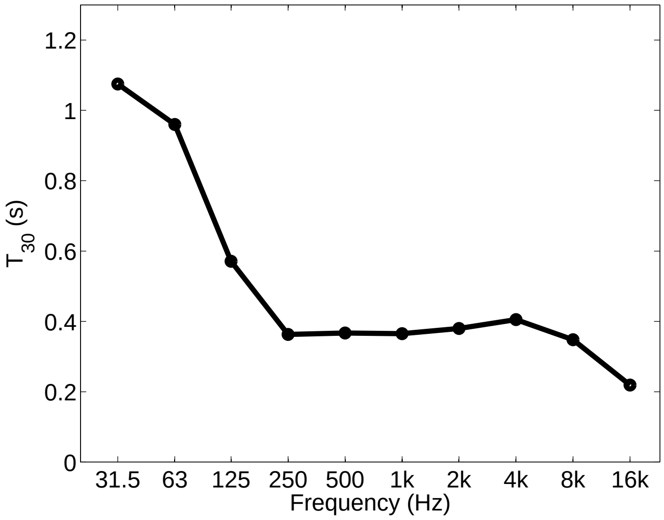


Figure 4.3: Measured reverberation time (s), T_{30}, of the test room.

4.2.2 Stimuli

The recording, playback, and live presentation of the stimuli all occurred in the same, acoustically normal room. The room was 4.7 m in length by 4.6 m in width by 3.4 m in height, forming a total volume of approximately 74 m^3. The room had a linoleum floor, decoupled double brick walls, and one window. The room's ceiling included a concrete upper surface and a suspended ceiling, which was suspended by approximately 0.6 m to a height of 2.8 m above the floor. The suspended ceiling was composed of perforated metal above which lay a thin layer of absorptive material. Additionally, there were four skylights in the ceiling. The room's reverberation time (T_{30}) between 32.5 Hz and 16 kHz is shown in Figure 4.3. The reverberation time measurement was made with source and receiver positions similar to that of the stimulus and listener in the listening tests.

Recordings of the test stimuli were made in a position typical of where the test subject would be seated, thereby maintaining a similar impression of both the direct and reflected sound as allowed for in particular by the binaural-recording technique. Figure 4.1 shows the recording setup with the acoustic manikin seated in the same position in which the test subjects later sat. The curtain, behind which the rods were dropped, is visible in the background. The headphones later used for playback and the length reporting apparatus are also shown in the photo. The microphone used for the simultaneous monophonic recordings was positioned over the head of the manikin at a distance of approximately 43 cm. Figure 4.4 shows simplified block diagrams of the three presentation methods.

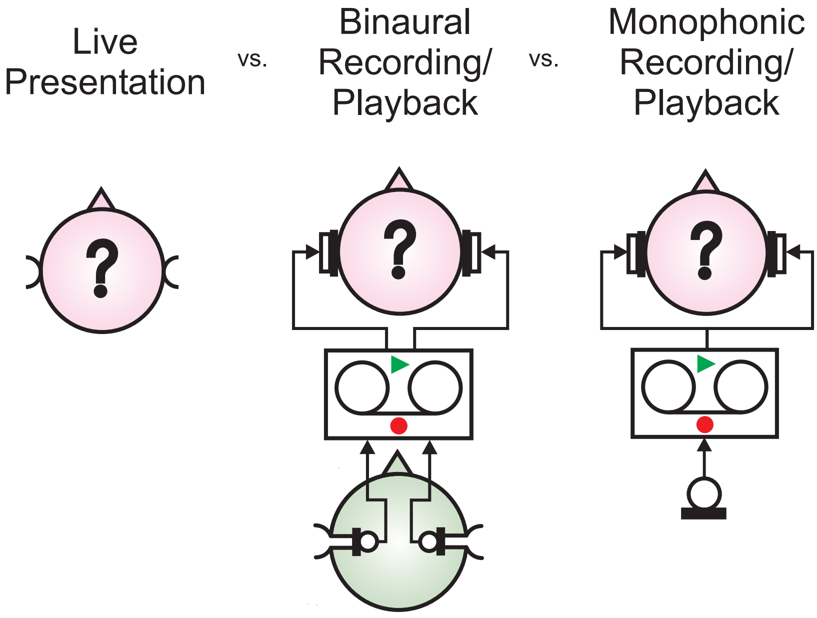


Figure 4.4: Simplified block diagrams of the three presentation methods compared. Microphones were placed in the ear canals at the positions of the eardrums of the acoustic manikin that was used for the binaural recordings

For recording, both of the manikin's microphones and the microphone used in the monophonic recordings were connected to a Brüel & Kjær NEXUS microphone signal-conditioning amplifier, Type 2690. The output channels of the microphone amplifier were connected to a Fostex D2424LV multi-channel hard disk recorder for 44.1 kHz, 24 bit recording. The hard disk recorder was kept in an adjacent room for noise isolation purposes. Recordings of the binaural and monophonic signals were made simultaneously, so that both a monophonic and binaural version existed of each rod drop. The monophonic and binaurally recorded stimuli later used in the test were therefore of identical sound events, only with a difference in the recording technique. All listeners heard both versions in the experiment.

The binaural stimuli were recorded using a Brüel & Kjær Type 4128 Head and Torso Simulator. This acoustic manikin contains microphones placed in the position of the eardrum. Playback for test subjects therefore required the influence of the ear canal in the manikin's head to be removed via equalization, as the listener would also have an ear canal through which the sound would travel when listening with headphones. The influence of the Stax SR Lambda Professional headphones with SRM-T1 headphone amplifier and their imperfect frequency response was also necessary to remove from the playback chain (Møller, 1992). This equalization along with that of removing the influence of the manikin's ear canals was effectively performed by fitting the manikin with the headphones that would later be used in the listening test and measuring a frequency response between the output of the microphones in the acoustic manikin and the input to the headphone amplifier. The inverse of the complex frequency response was then used to equalize the binaurally recorded stimuli prior to playback.

The monophonic stimuli were recorded using a Brüel & Kjær Type 4165 microphone with Type 2619 preamplifier. This microphone should have an approximately flat frequency response in the range of frequencies relevant to this stimulus. For the equalization of the monophonic recordings, it was therefore desired to be able to remove the influence of the remaining imperfect frequency response of the headphone playback system. The headphones were fitted on a Brüel & Kjær Type 4153, IEC 318 compliant, artificial ear for this purpose, and the frequency response measured. The frequency response was used to equalize the monophonic recordings in a similar manner as was used for equalizing the binaural recordings.

Participants were told that they would "wear headphones" for parts of the experiment, but it was intentionally left unstated that they would be listening to recordings. It was also left unstated where the sounds would come from or how they would be generated. A copy of the instructions of the listening test can be found in Appendix B. Open headphones, which do not seal off the ears from the outside world and which generally help to avoid occlusion, were used for playback in order to make the presence of the headphones as transparent as possible to the wearer. Loudspeakers mounted on the wall in front of the test subject may have added to the ambiguity of possible sound sources. Recordings were played back from a PC with a Hoontech SoundTrack Audio DSP24 sound card. The digital optical output was routed to a Digital Audio Denmark ADDA 2402 digital-to-analog (D/A) converter of which the output was connected to the headphone preamplifier. For noise isolation purposes, the computer with sound card and the D/A converter were kept in an adjacent room.

Playback sound levels were calibrated to closely match those of the live case. Calibration tones of a known level produced by a Brüel & Kjær Type 4230 sound level calibrator were recorded along with the original stimuli. The recorded (and equalized along with the stimuli) calibration tones were played back through the headphones into the manikin's ears from which the microphone output levels were measured. The digital signal levels in combination with the headphone amplifier volume were then adjusted to produce the appropriate levels. The level of the diotically-presented monophonic signal was later required to be further subjectively adjusted to be of equal perceived loudness as the live stimuli.

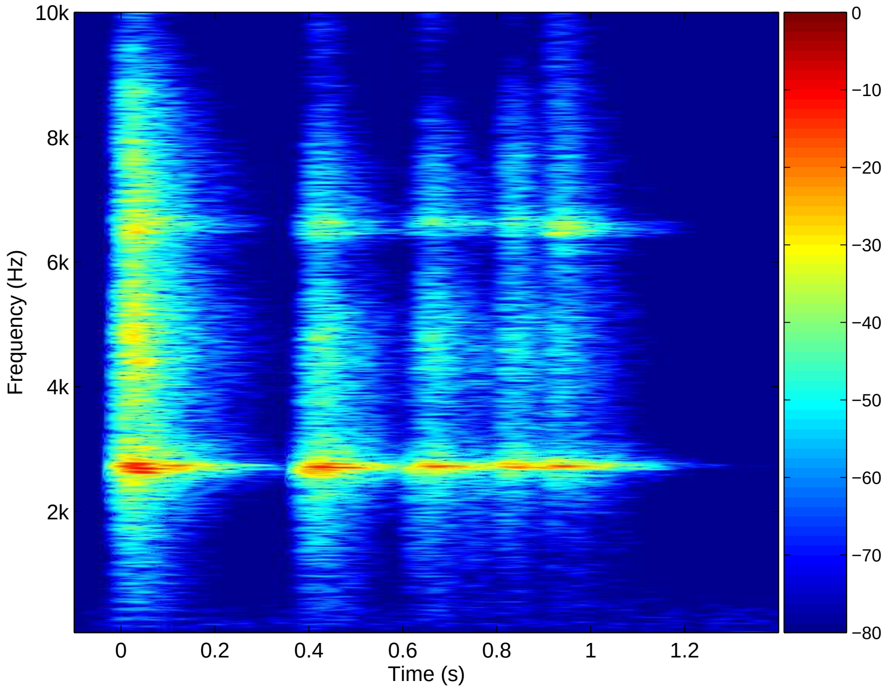


Figure 4.5: Spectrogram of the 15 cm long dowel dropped on the linoleum floor. Maximum levels are shown in red and minimum levels are shown in dark blue. The color decibel scale (relative) is matched to that of Figures 4.6 and 4.7.

Signal amplitude versus time plots and spectrograms for typical versions of the eight stimuli were produced with MATLAB and are shown in Appendix B. A few examples are given here. The full recordings used in the listening experiment were 2 s in length, but they have been truncated for presentation purposes.

A spectrogram of a 15 cm long dowel dropped on the floor of the test room is shown in Figure 4.5. Red indicates the highest intensity while dark blue indicates the lowest intensity. The initial impact occurs at about 0 s, followed by a second major impact at approximately 0.4 s as can be seen by the abrupt change across many frequencies at a single time. Other impacts follow these two initial bounces. It can also be seen in the figure that there are resonant frequencies (horizontal lines) at which the signal is strongest. For example, there appear to be resonances at approximately 2.8 kHz and 6.6 kHz.

A spectrogram of a 30 cm long dowel dropped on the floor of the test room is shown in Figure 4.6. Like in the previous figure, the initial impact occurs at about 0 s, followed by a series of other impacts. Compared to Figure 4.5, it appears that the lowest resonant frequencies have been shifted downward and there are

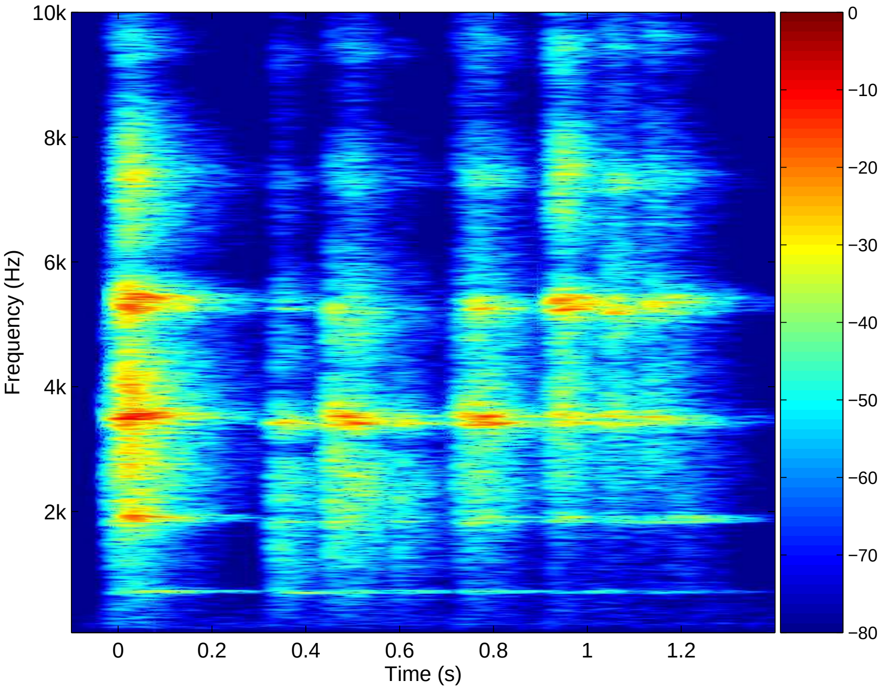


Figure 4.6: Spectrogram of a 30 cm long dowel dropped on the linoleum floor. Maximum levels are shown in red and minimum levels are shown in dark blue. The color decibel scale (relative) is matched to that of Figures 4.5 and 4.7.

more identifiable resonant frequencies. Visible resonances exist at approximately 600 Hz, 2 kHz, 3.6 kHz, 5.5 kHz, and 7.5 kHz. This downward shift in frequency can be heard when listening to the 30 cm long dowel being dropped on the floor after the 15 cm long dowel. The pitch of the 30 cm long dowel is lower than that of the 15 cm long dowel. Note that this is not to say that pitch is necessarily responsible for size perception, but simply that one can hear differences between the pitches of the signals. Fitch (1997) suggested that fundamental frequency is a poor indicator of vocal tract length and body size, so there is reason to suspect that there may be multiple cues suggestive of size information in general.

For even longer rods, the resonances are shifted even lower and become much more closely spaced. Figure 4.7 shows a spectrogram of the 120 cm long dowel being dropped onto the linoleum floor. Many more resonant frequencies have been excited in this case. One notices when listening that the pitch of the sound resulting from this dowel being dropped onto the floor is lower than that produced by the other dowels.

In addition to the resonant frequencies, it is apparent from the figures that there

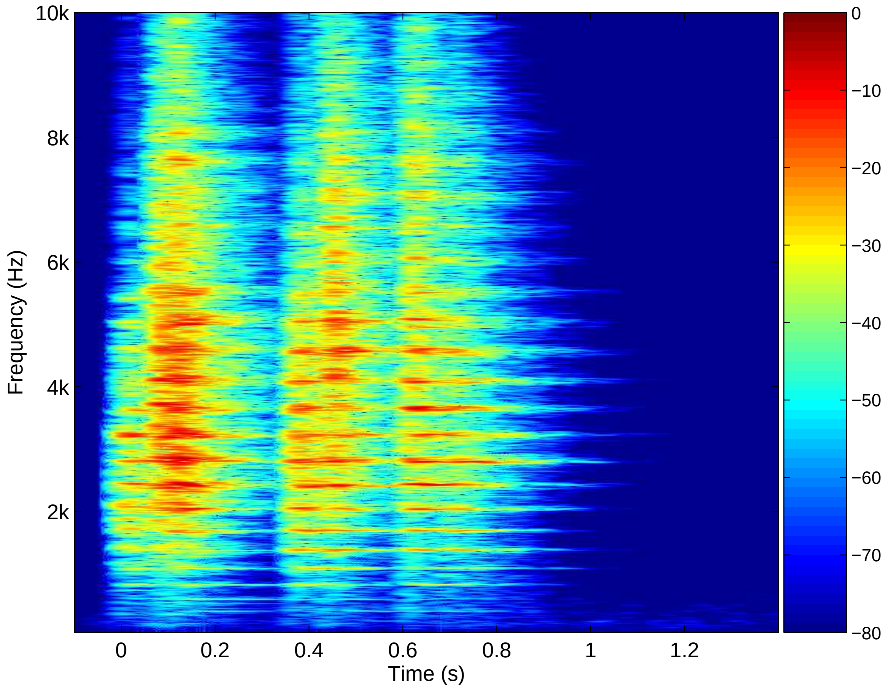


Figure 4.7: Spectrogram of a 120 cm long dowel dropped on the linoleum floor. Maximum levels are shown in red and minimum levels are shown in dark blue. The color decibel scale (relative) is matched to that of Figures 4.5 and 4.6.

Table 4.1: Mean perceived rod lengths versus actual rod length for each presentation method. Standard deviations are shown in parenthesis.

Actual Length (cm)	Mean Perceived Length (cm)		
	Live	Binaural	Monophonic
15	12.3 (4.2)	13.0 (3.7)	12.1 (3.6)
30	28.1 (6.5)	29.9 (7.1)	30.4 (9.9)
45	46.5 (9.8)	54.1 (12.4)	57.0 (19.7)
60	65.3 (15.7)	64.4 (9.1)	67.5 (14.3)
75	75.4 (7.9)	84.9 (13.9)	81.2 (19.0)
90	90.0 (14.5)	91.6 (16.1)	88.4 (17.9)
105	102.8 (14.1)	108.7 (12.1)	97.2 (20.0)
120	114.5 (11.6)	113.4 (14.0)	111.4 (18.1)

is a broad range of frequencies excited. It is not shown here, but there is also energy in the signals up to 20 kHz, though it is weaker than that present below 10 kHz. Examples of spectrograms resulting from the sound of rods dropped on a softer surface will be presented later in Chapter 5 and in Appendix C. Comparing the spectrograms of similar wooden rods dropped on two different surfaces demonstrates that the rods dropped onto a hard surface result in spectrograms with more high frequency energy. This tendency has also been previously observed by Ishibashi & Preis (2005b). It is possible that listeners are able to make use of information above 10 kHz, and the reader should not assume that this information is irrelevant because it is not presented here.

4.3 Results

An average of all response data for each presentation method indicates very accurate estimations of rod lengths for each of the presentation methods. Figure 4.8 shows that the mean length estimates for live presentation are generally closer to perfect performance (the diagonal line); however, the mean lengths for each method are not consistently statistically different at a 95% confidence level according to a two-way analysis of variance ($F_{2,14} = 1.88$, $MSE = 20.09$, $p = 0.19$). The data from Figure 4.8 are shown in Table 4.1 along with standard deviations in parenthesis after each mean perceived length.

A regression of perceived length onto actual length was calculated for the mean data for each of the presentation methods. The results are shown in Table 4.2. The r^2 values are very high for all presentation methods. The slopes are very near

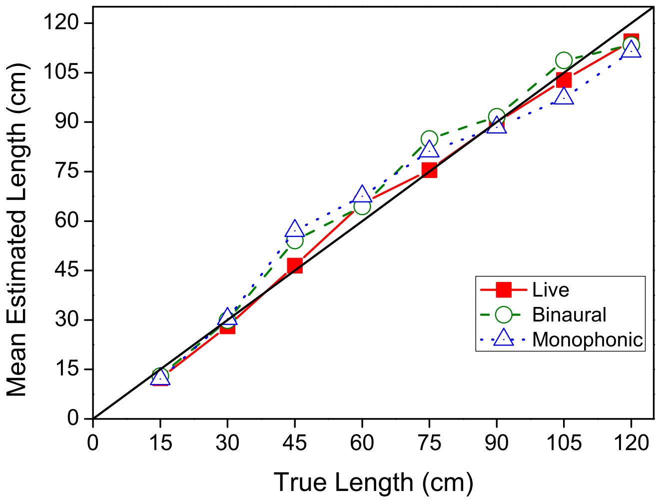


Figure 4.8: Mean length estimates for live (squares), binaural (circles), and monophonic (triangles) presentation methods. The solid diagonal line represents a perfect match between true length and mean estimated length.

Table 4.2: Results of a linear regression for mean perceived length onto actual length.

Presentation Method	r^2	Slope	Intercept (cm)
Live	0.99	0.98	0.97
Binaural	0.98	0.98	4.1
Monophonic	0.96	0.90	7.2

1.0 for the live and binaural presentation methods. Regressions were additionally calculated for the data of each individual subject for each method. The resulting r^2 values were between 0.83 and 0.99, with no obvious differences between the results for each presentation method. Regressions were significant for mean data ($p < 0.0001$) and for all individual subject data ($p < 0.01$).

While no notable differences were found in the group average data, inspection of the errors in the individual responses for each method reveals some differences between presentation methods. Figure 4.9 compares the mean percent error magnitudes (unsigned errors) for each of the presentation techniques. In order to calculate these mean error magnitudes, the mean estimated length for each rod was first calculated for each subject's five repetitions. The unsigned error magnitudes

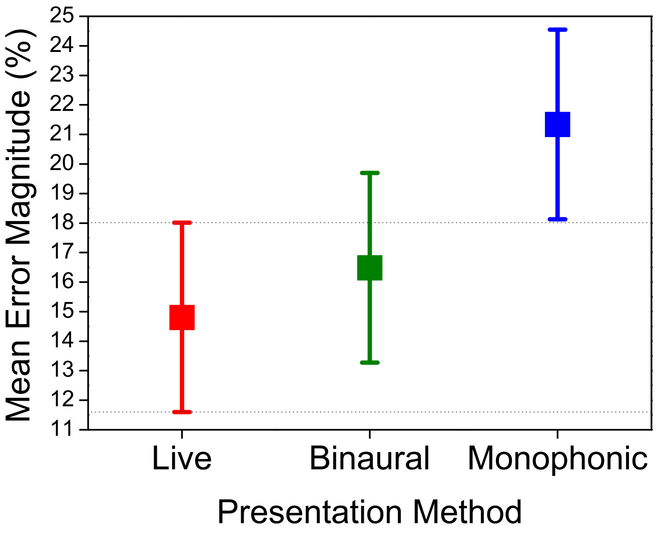


Figure 4.9: Mean length estimation percent-error magnitudes (squares) for each presentation method with multiple-comparison 95% confidence intervals shown.

of these estimations were then averaged across all subjects and all rod lengths, for each presentation method. When compared using multiple-comparison 95% confidence intervals resulting from an analysis of variance, it was found that subjects made larger errors on average when performing the task from monophonic recordings than from live presentation ($F_{2,189} = 3.07$, $MSE = 736.05$, $p < 0.05$). Subject performance when listening to binaural recordings was not found to be worse than when listening to live presentation (t-test: $p = 0.48$), but it was also not found to be better than when listening to monophonic recordings (t-test: $p = 0.10$).

In the case of the binaurally recorded stimuli, subjects were often unaware that they were listening to recordings during the listening test. Many subjects apparently believed they were wearing headphones for some other purpose, indicated by questions like, “Is it okay that the headphones aren’t working?” This was not the case for the monophonically recorded stimuli, for which it was generally obvious to the subjects that they were listening to reproductions. It is apparent from the subjective comments that there are properties of the sound, likely the spatial ones, present in the binaural recording and presentation technique that are important for maintaining realism.

Comments from test subjects indicated that some believed they were able to recognize rods being redropped in the test, but their stated guesses at how many rods were being used were always lower than the true number. Subjects mentioned that, as the test progressed, they might have been simply responding with length estimates that matched what they remembered previously reporting for a rod that they recognized. This provided reason to suspect possible training effects. However, comparisons of the data indicated no statistically significant improvement in accuracy, at a 95% confidence level, between the first and last sessions.

4.4 Discussion

The results of this experiment have confirmed the observations and hypothesis of VanDerveer (1979) that the use of an inappropriate recording and presentation method may falsely degrade test subject performance in a listening task. In the listening task described by VanDerveer, subjects were required to identify a variety of different sound events in a protocol study. The results of the present experiment also complement the findings of Møller et al. (1999), who found that the abilities of people to localize human speech were worse when subjects listened through an intermediate presentation method instead of listening directly to the stimuli. These combined results emphasize that recordings can degrade the perceptual fidelity of a sound event for a variety of sound events.

The mean length estimation accuracy for stimuli that were presented live has confirmed the results originally reported by Carello et al. (1998) that listeners are able to perceive the length of wooden dowels dropped onto a hard surface. While a relatively small number of test subjects participated in the present experiment and that of Carello et al., the similarity between the results suggests that this was not a problem for either case. Confidence in the results of each experiment is strengthened through the combined results.

While the test conditions, test stimuli, and analysis techniques are slightly different between the present experiment and that of Carello et al., the results appear to be of comparable accuracy, if not more accurate. The compression effect seen by Carello et al., in which the longest rods were reported by subjects to be somewhat shorter than they actually were, was not observable to the same degree in the present experiment. In the live condition, the mean perceived length of the longest rod was only slightly less than the actual length of the rod (actual: 120 cm; perceived: 114.5 cm). Furthermore, the mean perceived length of the shortest rod was slightly less than its actual length (actual: 15 cm; perceived: 12.3 cm). From these observations, it appears that the subjects made appropriate use of the length-reporting device and were not obviously restricted by the physical limits of the device.

The finding of this experiment that subjects did not improve in accuracy as the test progressed is consistent with that of Wagman (2003a). As described in Chapter 2, Wagman found that accuracy only increased when subjects were given visual feedback concerning their responses. The author also found that while accuracy did not increase without feedback, the consistency of the subjects' responses did increase. This is in agreement with subjective comments made by test subjects in the present investigation.

The exact results from this experiment and others likely depend on the specific dowels used in the experiment (e.g., their degree of homogeneity), the specific surface on which the dowels were dropped, the height from which the dowels were dropped, the angle at which the dowels were dropped, the position (relative to the subject) at which the rods were dropped, the way in which length was reported, and other factors. For example, if the subjects had been asked to perform this same task with the dowels dropped in front of them instead of behind them, the results may be better. While the precise results are therefore not directly comparable to other experiments, they should be comparable within the present experiment. Comparisons between presentation methods as made in this experiment should therefore be safe.

The finding that binaural recordings were not found to produce worse estimates than live presentation, but simultaneously were also not found to produce results that were better than monophonic presentation, is somewhat ambiguous. While it cannot be claimed that the binaural recordings produced more accurate length estimates than the presentation of monophonic recordings, it can also not be claimed that they produced less accurate length estimates than live presentation, which *was* found to be more accurate than the presentation of monophonic recordings. One could argue that, because of this, there is no reason to choose binaural recordings over monophonic recordings, but on the other hand, binaural recordings do appear to represent live events more accurately than monophonic recordings. For this reason, properly equalized binaurally recorded stimuli appear to be a better choice. Further tests with a greater number of test subjects may permit stronger claims to be made on this issue.

The most obvious difference between the binaural recording and playback technique and the monophonic recording and playback technique, as used in these experiments, is that the binaural technique maintains more spatial information. In theory, the binaural recording technique can allow a sound field to be perfectly reproduced, for an immobile test subject, if the original sound field is recorded and played back properly (Møller, 1992). However, some method limitations and deficiencies that existed in this experiment are worth noting when comparing presentation techniques. Sources of error between live presentation and the presentation of binaural recordings include the non-individualized head-related transfer functions and the fact that the reproduced acoustic environment will not stay in a fixed position if a listener turns his or her head in the environment. The result will be degradation in the accuracy of the sound field present at the test subjects' ears when listening to the recordings. The sound field will not be received in the same way the listener is used to in his or her everyday life. The influence of these issues for this experiment is not precisely known.

The monophonic recordings suffer from the above problems and an additional problem of great significance: the fact that the technique does not maintain any of the spatial information contained in the original stimulus. By its nature, the fact that the sound is recorded from only a single position, the angles of incidence of all sound waves, both direct and reflected, are forever lost. The monophonic technique's deficiencies that have been mentioned here, and which were further described in Subsection 2.3.1, are likely responsible for the differences in performance that were found in this experiment.

The major differences between live presentation and the presentation of either binaural or monophonic recordings is therefore that spatial cues are not perfectly represented unless the listener's head has been immobilized. This deficiency is much worse for the case of presenting monophonic recordings in which all of the spatial information and much of the information concerning position is absent. If monophonic recordings can otherwise be considered to be accurate representations of live sounds, for example that the absolute and relative levels of the frequencies present in the stimuli are adequately reproduced, then it is the missing spatial information in the monophonic recordings that can be blamed for their worse ability to represent the stimuli. This would suggest that the results of this experiment may indicate that spatial cues are important to the perception of object length; however, this hypothesis should be verified. An experiment that showed that binaural recordings were more accurate at representing a stimulus than monophonic recordings would be better proof of this. Evidence that is even more convincing could be generated by an experiment in which spatial cues were controlled more stringently. For example, by presenting the signal recorded from a single "ear" of an acoustic manikin to both ears of a test subject, a great deal of spatial information could be removed from a signal. This idea will be discussed more in Section 6.4 of the General Discussion of this thesis.

Monophonic recordings are clearly to be avoided if stimuli are to be represented as accurately as in the real world. However, if for some reason monophonic recordings must be used, there is an alternative way of presenting them that may provide a more ecologically valid experience for the test subject. Instead of asking test subjects to use headphones to listen to monophonic recordings, anechoic recordings of stimuli could be made and then played back to test subjects via a loudspeaker in a normal room. This idea was discussed in Subsection 2.3.1. Its primary benefits are that the subjects do not need to wear headphones and that the sound will appear to come from outside of the subjects' heads, not from within their heads as is often experienced when wearing headphones. Many spatial cues that would normally have been present in binaural recordings are still lost when using this method, but at least the sound will appear to come from a point external to the heads of the subjects. This may have worthwhile benefits. It should be noted that such a setup would work best with dry recordings (e.g., recordings made in an anechoic environment) so that only the room in which the sound is played back is responsible for adding reverberation to the stimuli.

It should be emphasized that the binaural stimuli used in this experiment were equalized to account for the imperfect frequency response of the recording and playback system. It is not certain that the results would have been the same without this step. It appears that only a few of the researchers doing ecologically motivated experiments with everyday sounds have taken care to account for imperfect frequency responses (e.g., Ishibashi & Preis, 2005b). The influence of this is unknown.

Spatially accurate presentation methods, such as live presentation, afford information on size. Whether auditory cues indicate an object to be of miniscule size or of large size, the cues provide information that specifies the nature of the sound event. Therefore, even objects of small size are potentially misrepresented by a monophonic recording. The ability of listeners to identify and extract information from sounds of many kinds is therefore best maintained when listeners are presented with stimuli in a realistic way.

Other than its effects on perceived object size, a presentation method that does not represent other spatial attributes of sounds as they would be perceived in the real world further discounts the perceptual validity of the stimuli being used. The choice of a presentation method that ignores the importance of spatial properties effectively eliminates ways in which information is transmitted to listeners. It has been described in previous chapters that the sizes of objects involved in sound events, the facing angle, the relative motion, and the relative distance of the sound events, all help to specify the nature of sound events. Monophonic recordings chip away at this information, thereby excluding many of the cues, which are all known to be perceivable, from the identity of the sound event.

Presumably, size information may be made available from everyday sounds other than impact sounds. It is likely that there may be size information in all kinds of vibrating objects, for example the other sound events shown in Figure 2.1 on page 17: scraping, rolling, and deformation. Aerodynamic and liquid sounds likely also carry some information about size, but this conclusion may perhaps not be drawn as easily as the suggestion that other events within the *vibrating objects* category may contain size information. Events classified as *vibrating objects* have more in common amongst themselves concerning how sound is generated and radiated from them, hence their being grouped together (Gaver, 1993c).

The large variation in sound level produced by the impact sounds of this experiment, in combination with the finite dynamic range of the recording equipment, meant that it was necessary to take great care in the recording process in order to avoid creating distorted recordings while simultaneously not introducing an audible noise floor. The recording process inevitably produced a few stimuli that were distorted due to clipping. These stimuli were necessarily excluded from the experiment, but there may have therefore inadvertently been introduced a bias towards reducing some of the natural variation in level that would occur with live presentation. In this case, the perception of length from recorded stimuli for the longer rods, which generally produce higher sound levels when dropped, may be effected. Fortunately, the results show no obvious degradation in test subject estimates of long rod lengths.

In order to further investigate the acoustic cues that are relevant to the perception of length, it would be interesting to artificially vary properties of the stimuli in much the same way that monophonic recordings have removed spatial information from the stimuli in this experiment. For example, one could investigate whether or not sound level is important for perceiving length by normalizing recorded stimuli such that maximum sound pressure level differences were removed. Others have noticed that normalizing stimuli according to their sound pressure levels have reduced the abilities of listeners to judge ecologically relevant properties such as speed, size, and weight (e.g., Houben, 2002; Grassi, 2002; Ishibashi & Preis, 2005a). These results suggest that it is important to present stimuli at their natural levels. Additionally, the results may have important implications for hearing aid compression systems. High compression ratios in combination with fast time constants, which together reduce natural sound pressure level differences, may similarly result in worse perceptual judgments. Experiments involving subjects listening to stimuli that have been compressed in various ways could provide further support for this hypothesis.

Further investigations of the acoustic cues important for perception may find it useful to use rods that are more homogeneous than the pine dowels used in this experiment. On one hand, it is interesting to see if listeners can accurately recognize the lengths of rods even with rods of varying homogeneity, but an experimenter who is interested in carefully controlling this aspect of the experiment may find it helpful to use other materials or other types of woods with which the homogeneity can be more stringently controlled.

While the length reporting device used in this experiment appears to have been a success, a possible downside of it was that the subjects were not in a fixed position for every trial. Therefore, when listening to recordings, their heads were not at the same position at which the binaural recordings had been made. The spatial representation of the sound field at the point at which they were listening was therefore not a perfect representation of the sound field that would have existed had recordings been made from the exact position of the subject. The recordings were made from a single, fixed position, at a single angle, whereas the test subjects listened from many other positions and with their heads at many other angles. This disadvantage was accepted knowingly, as it was preferred to live with its consequences instead of restricting the motion of the test subjects' heads using a bite bar or other head-restraint device.

A head tracking and stimuli synthesis system could have alternatively been used to produce stimuli for the subject in a way that adequately represented the general direction of the source, but as mentioned in Subsection 2.3.1 on page 58, such a system would still not be capable of representing other spatial qualities of the sounds. It is therefore unclear that a head-tracking and live stimuli synthesis system would be of any advantage.

One way in which the position of the subject could have been better maintained would have been to not require the subject to move around in order to use the length-reporting device. This solution is attractive from the perspective of being able to produce a more accurate sound field at the ears of the listener, but it may have actually been an advantage that subjects were forced to move around and be physically involved in reporting length. By requiring the subjects to move, they have become more directly involved in their responses. Active response tasks have been shown by others to improve the accuracy of perceptual judgments (e.g., Bootsma, 1989; Rieser et al., 1990; Loomis et al., 1992; Ashmead et al., 1995; Neuhoff, 2001a; Witt et al., 2005; Russell & Schneider, 2006), and it is possible that these effects may have contributed to the accuracy of the results observed here.

The fact that the presentation of monophonic recordings resulted in worse length estimation performance than did the presentation of live stimuli may have far-reaching consequences for auditory research, which is dominated by monophonically recorded or synthesized stimuli. VanDerveer (1979) wrote, "For studying the processes involved in hearing, recognizing, and remembering events, recordings must be made with care, and checked for their intelligibility. The ears of one or more listeners provide the best test.” When using any kind of recordings in place of live presentation, great care should be taken to avoid making sacrifices that may degrade the listening experience and potentially compromise ecological validity.

The results from this experiment indicate that subjects make larger length estimation errors, in an auditory length estimation task, when listening to monophonic recordings than when listening to stimuli presented live. Simultaneously, the monophonic technique did not provide the same sense of veracity that the binaural technique produced. It appears that equalized binaural recordings offer a better solution when recordings are chosen to be used, if for no other reason than the subjective realism experienced when using such recordings. Monophonically recorded stimuli should be used with caution.

As was done by Cooper et al. (2001) and Wagman et al. (2005a) (see also Wagman et al., 2005b), further investigations are needed in which subjects are asked to judge the lengths of rods at the same time that other properties of the impact events are varied. Such tests can provide stronger and even more ecologically relevant data concerning the abilities of people to perceive length. Stronger proof of the ability to hear the lengths of rods would be provided by an investigation in which listeners were asked to judge the lengths of rods at the same time that other stimulus parameters such as material and diameter were also varied. Chapter 5 describes an experiment that investigates this issue and other issues related to the perception of impact sounds.

Chapter 5

The Influence of Hearing Impairment and Hearing Aids on Impact Sound Perception

> *Laboratory studies concerning pitch perception, loudness, and so forth are as distant from these sorts of questions as studies of people's perception of letter shape and spacing would be from understanding reading comprehension.*
>
> -William W. Gaver (1993c)

5.1 Introduction

The importance of studying auditory perception, not only auditory sensation, is slowly being recognized among the hearing-research community (e.g., Plomp, 2002; Neuhoff, 2004b). Through classical psychoacoustical tests, a great deal has been learned about the physiological functionality of the human auditory system, but by having ignored practical uses of the system, little has been made clear about how it works, and even what it is used for, in the everyday world (Rosenblum, 2004). For some reason many people think of events in the world as visual, but events often have consequences for many perceptual systems, including the auditory system (Jenkins, 1985). Normal-hearing people continuously acquire a great deal of information about their environments from the sounds entering their ears (Gaver, 1993c). For situations such as leaves blowing in the wind, a rubber ball bouncing on a hardwood floor, rain falling on a tin roof, or a colleague typing at her desk, information about events and the objects involved in those events is constantly provided to listeners in their environments.

Sounds produced by impacting objects are one way in which information about the environment is carried to people (Gaver, 1993c). Impact sound events involving foot steps (Li et al., 1991), hands clapping (Repp, 1987), falling objects (Warren, Jr. & Verbrugge, 1984; Carello et al., 1998; Ishibashi & Preis, 2005b), struck objects (Gaver, 1988; Freed, 1990; Kunkler-Peck & Turvey, 2000; Giordano & McAdams, 2006), among others, have been shown to be acoustically informative to normal-hearing listeners about the events and about the properties of the objects involved in the events. It has even been demonstrated that impact sounds can convey information underwater (Tucker, 2003). Properties of impact events that may be transmitted by sound include the materials of the objects involved in the impact, the sizes of the objects, the shapes of the objects, the intensity of the impact, the location of the impact including its distance and direction from the listener, and many more.

Multiple properties of the objects involved in sound-producing events are transmitted to people in their daily lives. The material, size, and height from which an object fell, are examples of properties that may be provided by an impact sound resulting from a dropped object. Listeners perceive these characteristics, properties of the source events themselves and not of the sounds, simultaneously. It therefore makes sense that when studying listeners' perceptions of sound producing events, such variables should be both varied simultaneously and questioned simultaneously. As is done in traditional psychoacoustic experiments, changing only one variable at a time and asking about only that variable may produce results that are of little ecological relevance. Listeners may detect changes in sensations without being able to identify their cause. Varying multiple parameters simultaneously and asking listeners to detect these changes seems to be a stronger test of whether the information is present and useable (Gaver, 1988).

A few of the properties of an acoustic signal that may be involved in carrying information to listeners are the overall sound pressure level, spectral shape, the temporal envelope of the signal, spatial information from the direct and reflected sound received by the listener, higher-order time-varying versions of these parameters, and others. Apparently, multiple signal properties could carry the information relevant for perception. One should be careful before assuming a single attribute is solely responsible for an eventual percept. Li et al. (1991) have made similar conjectures. As an example, spectral attributes may not be the only attributes informative about the length of a rod dropped on a hard surface, as some may be quick to assume. Many acoustical characteristics could be informative about the length of a dropped rod:

- Sound pressure level - louder could indicate greater length;
- Temporal envelope - long rods may bounce differently than short rods;
- Spectral attributes - dominant resonances located at a high frequencies may indicate short rods;
- Spatial attributes - the sound from a shorter rod may appear to come from a more narrow area.

However, all of these acoustical characteristics could also be modified by changes to other physical properties:

- Sound pressure level - may be lower if the rod is dropped from a lower height or from a further distance from the receiver;
- Temporal envelope - may vary depending on the height of drop;
- Spectral attributes - may vary if the diameter of the rod is varied;
- Spatial attributes - may vary depending on the orientation of the rod.

It is interesting to note that for these acoustical characteristics, a single characteristic alone could not be solely responsible for informing a listener about length because these characteristics may also change when other physical attributes of the stimulus changes. Therefore, for an everyday listening task, multiple potentially informing acoustical characteristics must be evaluated by the listener together as a whole, simultaneously, if successful identification is to take place. While some acoustical cues may provide redundancy in the information being transmitted, it is difficult to pinpoint one cue that is more important than the others are. Attempting to do so may be counterproductive. Handel (1989) stated that, "The strength of the whole emerges from the connection among the parts."

It has been demonstrated that material is one property of objects involved in impacts that is detectable by listeners, solely on the basis of sound. When listening to live presentation of stimuli or to recordings of authentic stimuli, normal-hearing listeners are nearly perfect at being able to categorize materials that are grossly different from one another – materials such as metal and wood (e.g., Kunkler-Peck & Turvey, 2000; Tucker, 2003; Giordano & McAdams, 2006). Size is another property of objects that can be perceived from impact sounds. It has been demonstrated that listeners are capable of ranking the sizes of objects involved in impact sounds, and that their estimates of size may often be of reasonably high accuracy (e.g., Carello et al., 2005; Grassi, 2005; Kirkwood, 2005c).

Knowledge of the material and size of objects helps guide the actions of listeners able to obtain it (VanDerveer, 1979; Gaver, 1988). The listener is, in daily life, an actor in the environment in which sound events occur, and his action is guided by these events (Turvey, 1986). Information about material and size, as with all information carried by everyday sounds, can be used by a listener to decide if action should be taken. Not only can decisions related directly to action be made, but also by being confident that no action is required, a degree of comfort may be provided.

For the case of impact sounds produced by fallen objects, it is also clear that listeners must have some idea of the height from which an object has fallen. As an example, imagine a steel bar being dropped from the top of a building that is undergoing construction. It is clear that there are differences between the sound generated in that case and the case of that same steel bar being dropped by a construction worker on the ground. Information concerning the height of a drop could be informative to listeners in that it would further help identify the acoustic event that has taken place. If an object has fallen from a high height, perhaps evasive action should be taken to prevent injury. If the cat has knocked over a glass of wine, it may also be of interest to attend to the situation. Researchers in laboratory environments (e.g., Ishibashi & Preis, 2005b) have manipulated drop height in their experiments involving impact sounds, but without asking participants about this height. However, it has been shown that normal-hearing listeners are able to take advantage of auditory period information in ecologically valid listening tasks (Warren, Jr. et al., 1987).

It has also been demonstrated by Warren, Jr. & Verbrugge (1984) that the sounds of initial and successive impacts, occurring when an object is dropped, are informative about whether the object breaks or bounces. Their work showed that temporal properties are informative of the differences between breaking and bouncing events. In another study, Grassi (2002) showed that the bounces resulting from a ball dropped on a surface contribute to listener's abilities to judge the size of the ball (for an explanation of how temporal properties may convey size information, see Grassi, 2005). When bounces after the initial impact were artificially removed from recorded stimuli, judgments of ball size became worse. Presumably, similar temporal properties may also be informative about the height from which the object is dropped. Figure 5.1 shows a caricature of a bouncing event. Two acoustic cues that may be informative about height are immediately clear: the intensity of the impacts and the temporal pattern of the impacts as the object bounces. The height of the drop, among other factors, will determine the intensity of the initial impact. The ball in Figure 5.1, which is dropped from the higher height, will produce a more intense sound, all other factors being equal. The height of each

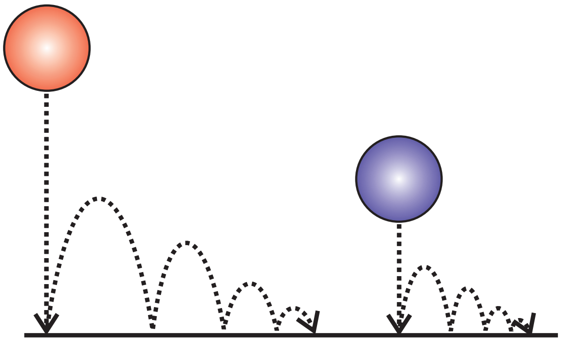


Figure 5.1: Conceptual diagram of a bouncing ball. The ball traces a pattern on a (moving) wall behind it as it bounces up and down until coming to rest. Ball on left is dropped from a high height and has a longer time between bounces than the ball on the right, which is dropped from a low height. [Figure inspired by Warren, Jr. & Verbrugge (1984) and Warren, Jr. et al. (1987)].

bounce, and therefore the time between bounces, depends on things such as the velocity of the object, the *material* of which the object is made, and the *material* of the surface on which the object is dropped.

Repeated bounces of an object dropped onto a surface can be seen when examining spectrograms of audio recordings made of the impact events. The spectrograms allow one to see how the short-term spectrum of the signal changes with time. Bounces show up as sudden changes of the spectrum with respect to time. Note that although it may be tempting to try to fully describe these signals by their spectrograms, it should be kept in mind that there may be other acoustic features which may not be readily, visually apparent from spectrograms. For example, spatial information that would be available to a normal-hearing listener is missing. Handel (1989) emphasized that it is important to evaluate whether the spectrogram is a relevant tool for use in understanding perceptual cues in sound. Nevertheless, spectrograms may provide at least *some* insight for an acoustic analysis of such events.

With the exception of experiments explicitly testing spatial hearing (for a review, see Blauert, 1996), spatial influence is one property of sounds that is very often ignored in psychoacoustic experiments. However, in daily life, the spatial relationship between objects and the environment is an integral part of how a person perceives and acts in the environment (Gibson, 1966, 1979). Therefore, it is evident that for promoting ecological validity, accurate spatial representations of signals are important to listening tasks. In testing auditory perception, it has been shown that the way in which listeners are presented with sounds has an influence on their perceptions of those sounds (e.g., Minnaar et al., 2001; Kirkwood, 2005c;

Çelik et al., 2005; Guastavino et al., 2005). In Chapter 4, experiments comparing various forms of stimulus presentation were compared for an everyday listening task. In this chapter, a new presentation tool is tested on both normal-hearing and hearing-impaired listeners: the hearing aid.

The hearing aid itself is essentially a recording and playback device, just as the two reproduction techniques described in Chapter 4, but the hearing aid attempts to make this recording and playback occur without delay. If hearing aids are used in a pair, this reproduction technique can be considered most similar to the binaural recording and playback technique used in the experiment described in Chapter 4. In or around each ear, a "recording" is made by a microphone. That sound is then processed and played back for the wearer at a position in the ear canal – not far from the position from which the original "recording" was made. However, unlike in the presentation methods previously considered, the processing stage in between the pick up and playback of the sound is intended to modify the sound for the wearer. More specifically, it is intended that the sound shall be modified in a way that *helps* the user (e.g., by amplifying it). In theory, this improvement should be based on the specific auditory requirements of the intended wearer. In practice, the user's needs are never fully understood and even the requirements that are understood may not be met.

Although some knowledge exists concerning the perception of everyday sounds by normal-hearing listeners, few researchers have conducted ecologically motivated experiments pertaining to how hearing-impaired listeners perceive sounds of non-speech and non-musical origin. It is not clear whether the hearing impaired are able to gather as much information from everyday sounds as are the normal-hearing population. Speech investigations excluded, few researchers have examined the abilities of hearing-impaired listeners to gather ecologically relevant information from their environments. Research of auditory problems has primarily focused on the physiological deficits associated with hearing impairment, and attempts to understand these deficits through psychoacoustic tests. While anecdotal evidence exists that hearing-impaired listeners, both with and without their hearing aids, may have trouble hearing properties of objects and events in their environments, few formal investigations have been done. Knowledge of how well hearing-impaired listeners can hear sound event properties compared to how well normal-hearing listeners hear the same properties could be of great benefit in understanding whether there is a need to help the hearing impaired with this task. Historically, most efforts at improving hearing for the hearing impaired have been aimed at improving speech intelligibility. Research has also been conducted to help insure that hearing aids are comfortable to listen to. It is not known whether the hearing aid processing done in an attempt to improve speech and comfort is a help or a hindrance to the perception of everyday sounds. Experiments to investigate this were conducted and are presented in this chapter. The listening tests were conducted using live presentation of the stimuli.

The experiment described in this chapter is an initial investigation of whether hearing-impaired listeners are as capable as normal-hearing listeners in hearing three ecologically relevant properties of impact sounds resulting from rods dropped on a floor: 1) the materials of the rods involved in the impacts, 2) the length of the rods, and 3) the heights from which the rods are dropped. All three of these properties, in addition to other physical properties of the stimuli, were varied simultaneously, and listeners were asked to describe their perception of the properties. It was hypothesized that the hearing-impaired subjects who participated in the test would make less accurate judgments of these properties than the normal-hearing subjects would. The results of listening tests are presented in which two subject groups, normal hearing and hearing impaired, have been tested with and without hearing aids. Testing both groups with and without hearing aids allows a test of the hypothesis that the hearing aids are beneficial, and not detrimental, in the perception of everyday sounds. The results are therefore informative about the abilities of normal-hearing listeners, about the abilities of hearing-impaired listeners, and about the influence of the hearing aids used in this study.

5.2 Methods

5.2.1 Subjects

Normal-Hearing Subjects

Sixteen normal-hearing subjects participated in the listening test. They ranged in age from 22 to 31 and had an average age of 27. They all had hearing thresholds better than 15 dB HL. Half of the subjects were male, and half were female. They were fitted bilaterally with behind-the-ear (BTE) hearing aids that had been programmed to provide a small amount of gain. Behind-the-ear hearing aids were chosen because they did not require custom earmolds to be made for the test subjects. The earhooks of the behind-the-ear hearing aids were connected to short, bent tubes that ran to foam ear inserts that were inserted into the test subjects' ear canals. Foam eartips from E·A·R Auditory Systems, which are normally intended for use with insert earphones, were used. The eartips have tubes running through them, allowing the sound produced by the hearing aid to pass through, but attenuating the direct sound, just as a common foam hearing protector would do. These foam eartips were used in place of what would normally be a custom earmold that would typically be made from silicone, acrylic, polyethylene, or vinyl (Dillon, 2001). The eartips used in this experiment are different from typical earmolds worn by hearing aid wearers in that they do not have vents. The subjects therefore experienced occlusion, most noticeable when they spoke. However, for this test, it was considered an advantage that the eartips attenuated the outside sound, as it was desired for the normal-hearing subjects to be listening primarily to the sound processed by the hearing aid.

The hearing aids worn by the normal-hearing subjects were the Syncro V2 BTE model manufactured by Oticon A/S. Hearing aids without a volume control were selected so that the level would not be adjustable during the tests. The programming of these hearing aids was done in conjunction with an audiologist from Oticon. To program the aids, the hearing aid fitting software "Genie" version 7.0, produced by Oticon, was used. This software is what is commonly used by hearing aid dispensers when programming Oticon-brand hearing aids for clients. Because the hearing aids are not intended to function without any amplification, the hearing aids were programmed to have a small amount of gain across the entire frequency range. A hearing loss was assumed, and the software prescribed a fit. The software automatically took into account characteristics of the device such as the specified diameter of the tubing used and the fact that there was no vent in the earmold.

A flat hearing loss of 55 dB HL was specified for the left and right-ear audiograms of the hypothetical client. This level was chosen in order to make sure that the hearing aid was active and to make sure that the sound reproduced by the hearing aid was dominant over any direct sound reaching the ears of the subjects. Many other input parameters were used by the software for prescribing the insertion gain and other settings of the hearing aid. It was specified in the software that the hypothetical subject was between the ages of 17 and 59, was a long-term hearing aid wearer, and that the subject had used nonlinear hearing aids in the past. The normal listening environment of the hypothetical wearer was set to "variable". It was specified that the hearing aids did not have vents in the ear molds and that the earhook of the hearing aid would dampen the sound by 9 dB. Tubing of 2 mm diameter was used between the hearing aid and the foam eartip, and this diameter was entered into the fitting software. An identity of "energetic" was chosen, the compression time setting was set to "faster", noise management was turned off, dynamic feedback control was turned off, and hearing aid directionality was set to "surround" (omnidirectional). These settings were chosen in order to produce a fit with fast compression time constants and with a minimum of other features that might alter the input signals. It was hoped that this fit would be an appropriate

Table 5.1: Gain (dB) prescribed for each frequency band and input level (“soft”,“speech”, or “loud”) of the hearing aids worn by normal-hearing test subjects. The maximum power output (“MPO”) level is also specified.

	Channel Center Frequency (Hz)							
Parameter	250	750	1.5k	2k	3k	4k	5k	7k
Maximum Output	99	101	101	101	101	101	99	98
Loud Input Level	8	10	8	9	10	11	11	10
Speech Input Level	9	12	10	13	13	11	11	10
Soft Input Level	21	26	27	25	25	24	22	21

compromise between being simple and with sound processing that would be easily understood from a technical perspective, and at the same time that it would be fairly representative of real-world hearing aid processing.

Finally, in order to flatten the default prescribed insertion gain across frequency, the gain settings for the channels with center frequencies of 250 Hz, 750 Hz, 4 kHz and 5 kHz were each increased by 1 dB. Both hearing aids were programmed identically. Final gain settings and the maximum power output (MPO) for each channel are shown in Table 5.1. Note that soft input sounds are given the most gain. The hearing aids were programmed to compress the dynamic range of the incoming sound. Normal-hearing people tend not to suffer from reduced dynamic ranges as do the hearing impaired, so compression was not strictly necessary for the subjects, but it is a typical way in which sound is processed in modern hearing aids and therefore was implemented for this test.

The hearing aids had eight frequency channels. In each frequency channel, there were two programmable compression ratios making up part of the multi-stage compression system. An example input/output curve for the hearing aids worn by the normal-hearing subjects is shown in Figure 5.2. The example shown is for the hearing aid channel with center frequency of 2 kHz. For all of the channels, the “speech” input sound pressure level and compressor knee point corresponded to 65 dB. “Loud” corresponded to an input level of 85 dB SPL. The “soft” level corresponded to between 20 and 45 dB (variable) for the three channels with center frequencies of 1.5 kHz and lower, and “soft” corresponded to 45 dB for the five channels with center frequencies of 2 kHz and higher. The center frequencies and compression ratios of each channel are shown in Table 5.2. A compression ratio of 1:1 applied to input sound levels below “soft” and to those between “loud” and the MPO. Attack and release times are shown in Table 5.3.

In order to make sure that the sound reproduced by the hearing aid was dominant over the direct sound, a test was done to compare the level of the direct sound

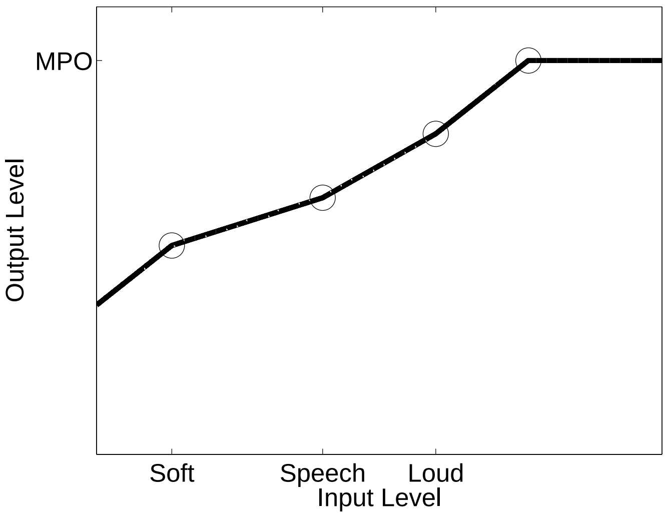


Figure 5.2: Sample input/output curve of the nonlinear hearing aid used by normal-hearing subjects. Compressor knee points are circled. Five distinct stages of the compression system can be seen: 1) below "soft" input levels, 2) between "soft" and "speech", 3) between "speech" and "loud", 4) between "loud" and the "MPO", and 5) above the "MPO". See text for a further description of the input levels corresponding to "soft", "speech", "loud", and "MPO".

Table 5.2: Compression ratios for hearing aid compression system. Compression ratios are shown for each frequency band of two ranges of input levels – "Speech to Loud" and "Soft to Speech".

	Channel Center Frequency (Hz)							
Compression Region	250	750	1.5k	2k	3k	4k	5k	7k
Speech to Loud	1.1:1	1.2:1	1.2:1	1.4:1	1.3:1	1.0:1	1.0:1	1.0:1
Soft to Speech	1.9:1	1.9:1	1.9:1	2.5:1	2.5:1	2.9:1	2.2:1	2.2:1

Table 5.3: Compression time constant settings for hearing aid compressors. Attack and release times are shown for each frequency band. Within a channel, the attack and release times were identical for all input levels.

Time	Channel Center Frequency (Hz)							
Constant	250	750	1.5k	2k	3k	4k	5k	7k
Attack	20 ms	20 ms	20 ms	10 ms	10 ms	5 ms	5 ms	5 ms
Release	80 ms	80 ms	80 ms	80 ms	80 ms	80 ms	80 ms	80 ms

reaching the ear to the level of the amplified sound. A white-noise noise source was set to produce 80 dB SPL. The sound pressure level, as measured at the eardrum position of an acoustic manikin, was measured in two conditions: 1) with the foam eartips and hearing aids in place but not turned on and 2) with the hearing aids in place and turned on. The results indicated a level difference of approximately 50 dB between the amplified sound produced by the hearing aid and the direct sound reaching the microphone, which should be plenty to insure that a test subject wearing hearing aids would be hearing the reproduced sound and not the direct sound. The comparatively low level of direct sound, which may leak through the blocked ear canals, would be masked by the sound produced by the hearing aid.

Hearing-Impaired Subjects

Twelve hearing-impaired subjects participated in the listening test. They ranged in age from 55 to 78 and had an average age of 63. Half of the subjects were male, and half were female. None of the subjects were born with their hearing losses. The severity of their losses covered a broad range, from mild to severe (according to categories defined in Gelfand, 2001). The only attempt to limit the subject group was to select only experienced hearing aid wearers who were believed to be suffering from sensorineural losses. Because sensorineural hearing loss is so widespread, and the fact that hearing aids are of limited use for those suffering from this type of hearing loss (Pickles, 1988), test subjects suffering from this type of hearing loss were chosen for the investigation. Subjects known to be suffering from conductive hearing losses, losses that result from abnormalities before the cochlea, were excluded.

Mean audiograms of both the hearing-impaired and normal-hearing subjects are shown in Figure 5.3. Figure 5.4 shows the individual audiograms for the hearing-impaired subjects. Some of the subjects in the experiment where chosen from a group of subjects who had participated in previous studies examining ski-slope hearing losses. For this reason, many of them have a similar audiogram configuration with relatively normal hearing at low frequencies followed by a sudden drop off at middle frequencies, ending with a flattening of the loss again at high frequencies.

The hearing-impaired subjects were all experienced hearing-aid users, each with at least a few years of experience. They therefore had an advantage over the normal-hearing subjects in the sense that the normal-hearing subjects were not experienced with wearing hearing aids. The hearing-impaired subjects wore a variety of hearing aid types, including open fittings (1 subject), behind-the-ear (4

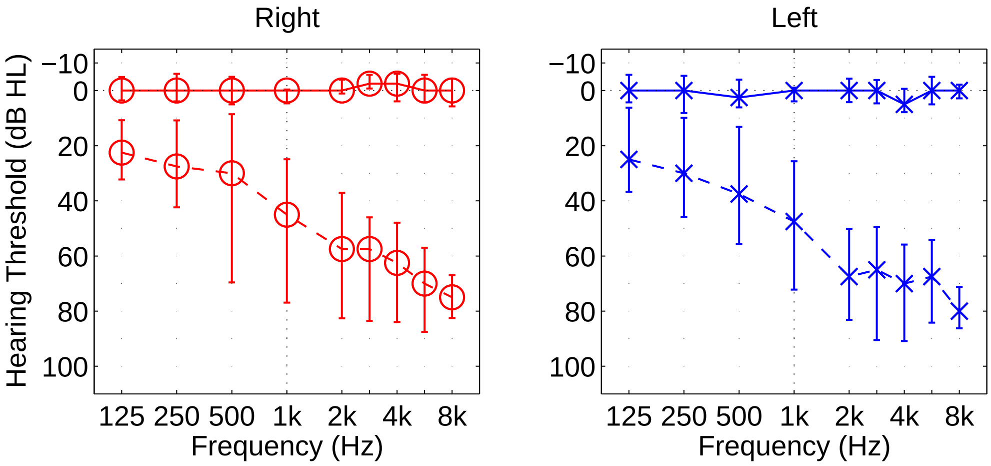


Figure 5.3: Mean audiograms of normal-hearing (solid lines) and hearing-impaired (dashed lines) subjects. Error bars indicate 20th and 80th percentiles. Plot aspect ratios in accordance with ISO 8253-1 (1989).

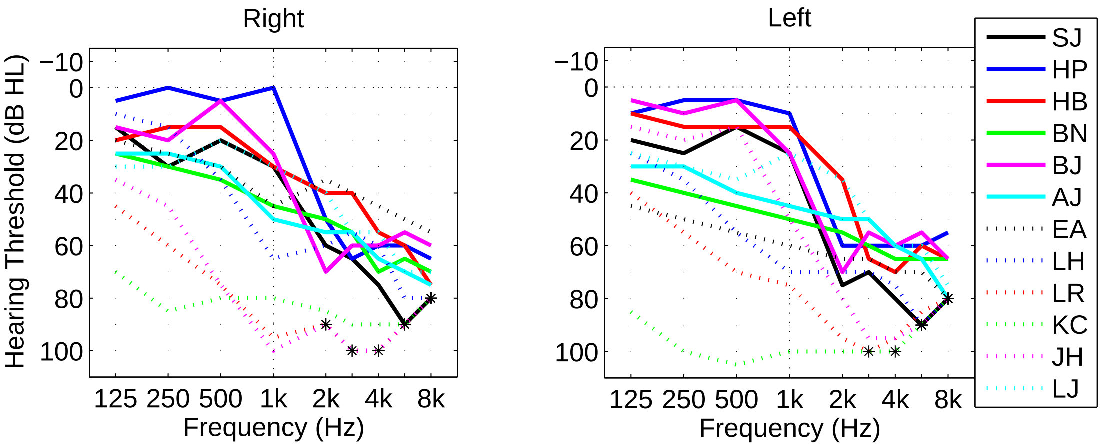


Figure 5.4: Individual audiograms of hearing-impaired test subjects. Asterisks indicate no response at the maximum output level tested. Plot aspect ratios in accordance with ISO 8253-1 (1989).

subjects), in-the-canal (5 subjects), and complete-in-the-canal aids (2 subjects). They were used to wearing them for at least four hours per day. All but one of the subjects, subject LR, had been using their current hearing aids for a few years, and all but one had bilateral fittings. Subject JH wore a hearing aid only in her left ear. Her right ear had a greater hearing loss for which she had not been fitted with a hearing aid.

The hearing-impaired subjects wore their personal hearing aids during the test. Their batteries were replaced as necessary before each test. For the listening test, the hearing-aid wearers were asked to use the mode of their hearing aid that they most often used. The option of providing hearing aids for the hearing-impaired test subjects was considered, but this would have required that the hearing aids be programmed for each subject. It was judged unlikely that a fit as appropriate as the one the subject was used to, with his or her personal hearing aids (a fit done by an audiologist), could be attained by the author (not an audiologist) performing a quick fit of the aids for the subject. Having the subject use his or her personal hearing aids had the advantage that all of the subjects were hopefully already acclimatized to using the hearing aids, if any acclimatization was ever necessary (for a review of acclimatization studies, see Turner, Humes, Bentler & Cox, 1996; Palmer, Nelson & Lindley, 1998). An additional advantage was that multiple types of hearing aids were represented during the test – hearing aids with realistic fits and which were used outside the laboratory as well.

Asking the hearing-impaired subjects to use their personal hearing aids during the listening test had the disadvantage that there were many parameters of the hearing aids that could not be controlled. These hearing aids could have had, and most likely did have settings which differed such as: compression attack and release times, compression thresholds, number of channels, noise control systems, directional-specific processing, gain settings prescribed by different prescription philosophies, acoustic-feedback managers that could have been activated by dropped rods, and many more. Most of these settings were unknown in this study. This could increase the amount of variability between the results of the test subjects, but in a manner hopefully representative of real-world differences between hearing aids and hearing aid fittings.

Hearing-impaired subjects were asked to complete a questionnaire concerning their satisfaction with their hearing aids. The Satisfaction with Amplification in Daily Life (SADL) questionnaire (Cox & Alexander, 1999, 2001) was chosen because of its moderately short length and its availability in Danish. The SADL questionnaire has 15 multiple-choice questions concerning satisfaction and results in a global score comprised of four subscales that cover the issues “positive effect”, “service and cost”, “negative features”, and “personal image”.

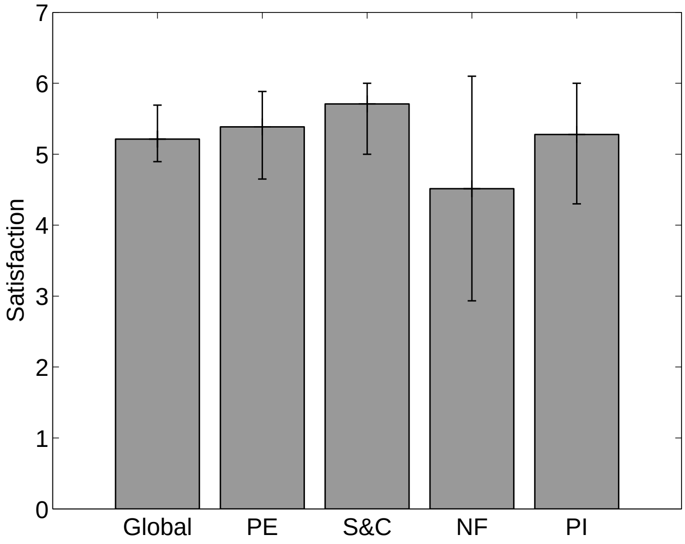


Figure 5.5: Mean scores in response to the SADL hearing-aid satisfaction questionnaire. Error bars indicate 20th and 80th percentiles. Individual categories are shown along the x-axis: Positive Effect (PE), Service & Cost (S&C), Negative Features (NF), and Positive Image (PI). Global Score summarizes all category results.

Results from the SADL questionnaire are shown in Figure 5.5. On a scale of 0 to 7 with 7 representing perfectly satisfied, the hearing-impaired test subjects had an average global score of 5.2. Global scores ranged between 4.3 and 6.7. The global and subscale satisfaction scores found for these subjects are similar to the norms described in Cox & Alexander (1999, 2001), indicating that the subjects who participated in this study had levels of satisfaction that are similar to the published norms. For the *service and cost* category, it should be kept in mind that none of the hearing aid wearers had to pay for their hearing aids, as they are provided free of charge to Danish citizens in need of them. As specified to be done by the authors of the questionnaire, the question of, “Does the cost of your hearing aids seem reasonable to you?” was therefore excluded from the analysis (Cox & Alexander, 1999).

In order to verify that the hearing aids worn by the hearing-impaired tests subjects were functioning, an Affinity hearing aid analyzer from Interacoustics A/S was used to perform acoustic measurements on the hearing aids. Measurements were made with a 2 cc coupler of basic input/output curves, frequency response curves (warble tone input), frequency response curves in response to 3-band, male, speech-shaped ICRA noise (Dreschler, Verschuure, Ludvigsen & Westermann, 2001), and temporal measurements designed to measure attack and release times. While a deep investigation of the results was not conducted, it was evident that the hearing aids were all operational.

5.2.2 Procedures and Stimuli

Subjects sat at a computer desk in an acoustically normal room. The room was the same used in the experiment described in Chapter 4. It was 4.7 m in length by 4.6 m in width by 3.4 m in height, forming a total volume of approximately 74 m^3. The room had a linoleum floor, decoupled double brick walls, and one window. The room's ceiling included a concrete upper surface and a suspended ceiling, which was suspended by approximately 0.6 m to a height of 2.8 m above the floor. The suspended ceiling was composed of perforated metal above which lay a thin layer of absorptive material. Additionally, there were four skylights in the ceiling. The room's reverberation time (T_{30}) between 32.5 Hz and 16 kHz is shown in Figure 4.3 on page 84. The reverberation time measurement was made with source and receiver positions similar to that of the stimulus and listener in the listening tests.

Subjects listened to rods being dropped on the floor behind them and answered questions concerning the characteristics of the rods and of the drops themselves. All experiments were conducted live – not from recordings. A diagram of the test setup is shown in Figure 5.6. The room had a linoleum floor, but the rods were dropped onto a thin floor mat at a distance of approximately 2 m behind the subject. The floor mat (essentially a thin rug with a thin rubber base) was used to reduce the peak sound levels to reasonable values and to prevent the rods from damaging the floor. Subjects were not told about the presence of the floor mat.

Twelve different rods were used as stimuli for the listening test. The twelve rods, dropped from two different heights, formed 24 unique stimuli. Three materials in two different diameters, each in two different lengths, were dropped from two different heights. Pilot tests were conducted to determine appropriate materials, lengths, and heights so that the test would be neither too difficult nor too easy for normal-hearing subjects. The length perception results from the experiment presented in Chapter 4 helped in selecting appropriate lengths. It was hoped that the results from the subjects would cover a range of judgment accuracies, and also allow room for differences between subjects and subject groups to be observable.

The rods were made of metal (anodized aluminum), wood (pine), and plastic (PVC-U electrical conduit pipe). All rods were presented in diameters of 16 and 25 mm.

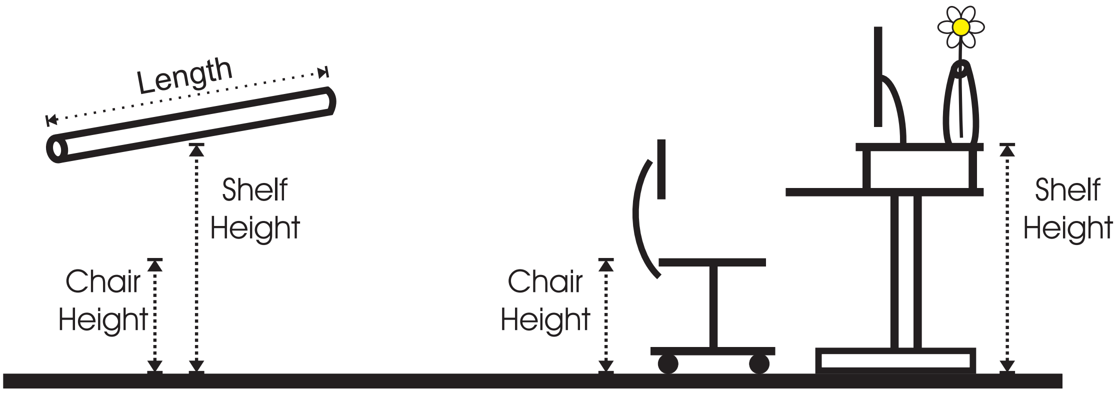


Figure 5.6: Diagram of listening test setup. Test subject sat at chair in front of desk (right). Rods of various lengths and materials were dropped from various heights onto the floor behind the subject (left). A visually opaque, but acoustically transparent curtain separated the subject and stimulus, but it is not shown here.

Rods of each material and diameter combination were cut to both 30 and 55 cm. The wooden rods were solid, but the metal and plastic rods were hollow. The walls of the 16 and 25 mm metal rods were 1 and 2 mm thick, respectively. The walls of the 16 and 25 mm plastic rods were approximately 1.4 and 1.7 mm thick. The masses of the larger diameter rods were therefore greater than the masses of the small diameter rods, adding yet another variable to which the subjects were exposed. The mass of each of the rods is shown in Table 5.4. A special fixture was used in an attempt to drop the rods in a consistent fashion on each trial. The rods were rolled off the edges of short plastic inclined planes. The rods were dropped at an angle of approximately 10° relative to the floor.

On each trial, subjects answered the following questions:

- "Of which material is the rod made?"
- "How long is the rod?"
- "From which height did the rod fall?"

For the question concerning from which material the rod was made, the subjects could choose "metal", "wood", or "plastic". In response to the length question, the subjects could choose from "short" or "long". The options in response to the question concerning from which height the rod was dropped were "shelf height" and "chair height". The response options are summarized in Table 5.5. Although it was varied (a fact which the subjects were *not* told), they were not asked about the diameter of the rods. Finally, if the test subjects did not hear the stimulus at all, they were asked to click on a button indicating so ("I didn't hear anything").

Table 5.4: Typical maximum fast-weighted sound pressure level (dB SPL) of each rod when dropped from shelf and chair heights. Mass of each rod is also shown.

Material	Thickness	Length	Mass	Shelf SPL	Chair SPL
Metal	25 mm	55 cm	203 g	107 dB	104 dB
Metal	25 mm	30 cm	111 g	100 dB	96 dB
Metal	16 mm	55 cm	61 g	101 dB	90 dB
Metal	16 mm	30 cm	35 g	99 dB	89 dB
Wood	25 mm	55 cm	145 g	100 dB	92 dB
Wood	25 mm	30 cm	74 g	91 dB	84 dB
Wood	16 mm	55 cm	57 g	94 dB	83 dB
Wood	16 mm	30 cm	30 g	88 dB	78 dB
Plastic	25 mm	55 cm	94 g	77 dB	75 dB
Plastic	25 mm	30 cm	52 g	75 dB	69 dB
Plastic	16 mm	55 cm	44 g	76 dB	74 dB
Plastic	16 mm	30 cm	27 g	73 dB	67 dB

Table 5.5: Response options for questions concerning the rod material, rod length, and height from which each rod was dropped. Properties are shown in bold with possible responses shown beneath each property.

Material	**Length**	**Drop Height**
Metal	Short	Shelf Height
Wood	Long	Chair Height
Plastic		

Subjects answered these questions via a computer interface. They responded to each trial by clicking radio buttons to answer each of the three questions on a computer screen. The questions were presented simultaneously, and the subjects had as long as they wished to respond. No time limit was put on the response due to expected large differences in skill levels at using the mouse. Subjects were told that they could answer the three questions in any order that they desired. The response menu was available in both English and Danish. Examples of the English and Danish response menus can be found in Appendix C on pages 191 and 193.

As mentioned, the heights were defined for the test subjects as chair height and shelf height. The desk at which the subjects sat was composed of a surface for a mouse to be placed and a raised shelf on which the display was placed (see Figure 5.6 on page 116). This shelf and the chair on which the subjects sat served as the references. These heights were 91 cm and 44.5 cm, respectively. The chair seat was wooden and nearly completely flat. A second, identical chair, adjusted to the same height as the chair on which the subjects sat, was placed next to the computer desk for visual reference. A vase containing a flower was placed on the shelf to give the subjects an idea of something that could fall from shelf height. These two heights were chosen for practical reasons and in an attempt to make the response options easily identifiable. The chair and shelf were concrete references in the subject's environment.

Colored tape, cut to the length of the short and long rods, was placed on the computer desk in order to show the subjects the two lengths from which they could choose. Blue tape was used for the short length, and red tape was used for the long length. Two examples of the lengths were placed on the desk. Both lengths were shown on the surface on which the mouse sat and both lengths were shown across the shelf ledge. This redundancy was simply done to allow the test subjects to more easily get to know from which lengths they could choose. It proved useful to have the length references across the front of the shelf, because the references on the desk were covered by the mouse pad for one test in which the subject used the mouse left-handed.

While the subjects had visual references for the lengths that they should judge and visual references for the heights that they should judge, no material references were shown to the test subject. It was assumed that all subjects knew well what metal, wood, and plastic were, and it was decided that no more information about the specific types of metal, wood, or plastic would be given.

Instructions for the listening test (see Appendix C, p. 190) were available to the subjects in both English and Danish. In an attempt to make it clear for the subjects what was meant by a "rod", the line drawing depicting one, as shown in Figure 5.6, was included for them in the instructions. They were explicitly told that the rods could be hollow or solid.

The test was composed of four sessions plus a brief training session. The training session consisted of six trials to make the subject aware of what kinds of sounds to expect and of how to use the response interface on the computer. None of the stimuli presented in the practice trials were identical to the stimuli presented later in the test, but they represented an approximate range of the sounds (i.e., materials, diameters, lengths, heights) that would later be heard. In each session, the subjects heard two repetitions of each of the 24 stimuli, for a total of 48 trials per session. In an attempt to reduce the influence of order effects, the sequence of rod drops was systematically randomized using Latin squares. Subjects completed two sessions with hearing aids and two without, therefore the subjects heard 96 trials per condition for a total of 192 trials in the experiment – each condition with a total of four repetitions. Half of the subjects completed the first two sessions of the listening test with hearing aids and half without. The two aided sessions were completed as a block, without removing the hearing aids, in order to avoid differences in the way the hearing aids were inserted each time. No feedback was given, and the subjects were not allowed to see the stimuli until after the test. The experiment lasted approximately one hour and fifteen minutes, including breaks, but varied quite a bit depending on the response times of the test subject.

A fan was run during the test in order to mask quiet sounds that might have inadvertently reached the room from the outside and to mask noises produced while the rods were handled. The fan was placed on the floor to the left of the test subjects at a distance of approximately 1.5 m. The sound pressure level at the position of the subject was approximately 50 dB. For the normal-hearing listeners wearing hearing aids, this level was amplified to approximately 70 dB. It was hoped that the fan may also help to provide a steady state condition for the hearing aid compression systems to return to between each trial. Had the room been perfectly quiet, subjects' hearing aids may have gone into very high gain modes, depending on how they were programmed. Inadvertent quiet sounds may have also caused the compression systems to flip between modes more frequently.

Acoustic Description of Stimuli

The typical maximum, fast-weighted sound pressure levels of each of the stimuli was measured with a Brüel & Kjær Head and Torso Simulator, Type 4128, connected to a Type 2636 measuring amplifier. The results are shown in Table 5.4. Note that while some of the maximum levels are quite high, these sounds occurred for only fractions of seconds, sporadically throughout the test. The equivalent continuous A-weighted sound pressure level, $L_{A,eq}$, was measured to be 84 dBA when measured over a period of one hour at the eardrum position of the left ear of the acoustic manikin. This test was performed with the manikin wearing hearing aids for half of the test, just as all test subjects did. An exposure level of 85 dBA for 8 h per day is considered safe (Moore, 2003).

Examples of all stimuli were recorded from the acoustic manikin's microphones connected to a Brüel & Kjær NEXUS microphone signal-conditioning amplifier, followed by a Sound Devices 744T hard disk recorder, which made 24 bit, 48 kHz recordings. The stimuli were recorded at the eardrum of the right ear of an acoustic manikin placed in the position at which a test subject's head would normally be located. Time signals and spectrograms for typical versions of all 24 stimuli were produced with MATLAB and are shown in Appendix C. A few examples are given

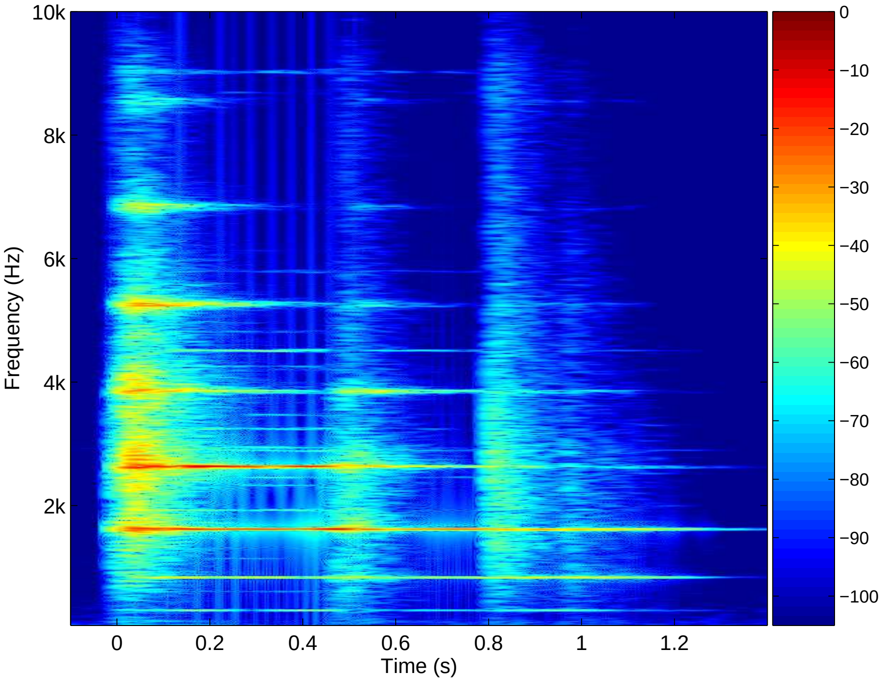


Figure 5.7: Spectrogram of a thin (16 mm diameter), long (55 cm), metal rod dropped from shelf height. Maximum levels are shown in red and minimum levels are shown in dark blue. The color decibel scale (relative) is matched to that of Figures 5.8 and 5.9.

here.

A spectrogram of a long (55 cm), thin (16 mm), metal rod, dropped from shelf height is shown in Figure 5.7. Red indicates the highest intensity while dark blue indicates the lowest intensity. The initial impact occurs at about 0 s, followed by a second major impact at approximately 0.45 s as can be seen by the abrupt change across many frequencies at a single time. Another bounce is easily observable just before 0.8 s, and finally there is another observable sudden change in the sound (impact) near 0.95 s. It can also be seen in the figure that there are resonant frequencies (horizontal lines) at which the signal is strongest, some of which resonate throughout the duration of the signal. For the metal rods used in this experiment, the resonances lasted up to approximately 2 s, although the full length has not been shown here.

An example spectrogram of a thick (25 cm diameter), long (55 cm), wood rod dropped from chair height is shown in Figure 5.8. It can be seen that the time between bounces is reduced compared to the previous example in which the rod

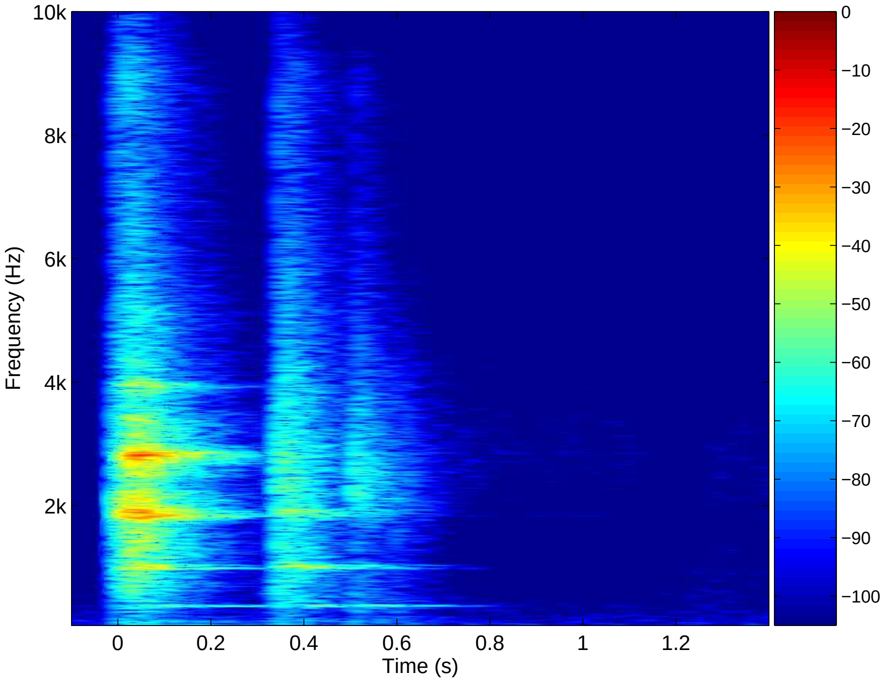


Figure 5.8: Spectrogram of a thick (25 mm diameter), long (55 cm), wood rod dropped from chair height. Maximum levels are shown in red and minimum levels are shown in dark blue. The color decibel scale (relative) is matched to that of Figures 5.7 and 5.9.

was dropped from shelf height. In this case, the first bounce after the initial impact occurs at approximately 0.35 s, followed by another bounce at approximately 0.5 s. The change in rod thickness, with respect to Figure 5.7 can be seen as a change in some of the resonant frequencies of the rod (horizontal lines). The resonant frequencies of this wooden rod have apparently lower quality (Q factor) values – that is, their peaks are not as sharp and do not fall off as quickly, with respect to spectral content, as those of the previously shown metal rod. Note that the resonant frequencies do not resonate for as long as the resonant frequencies of the metal rods did. The last visible sign of reverberation ends at just after 0.8 s for this plot, and there is a tendency for the resonances to somewhat die out between bounces.

A spectrogram of a plastic rod is shown in Figure 5.9. The rod is thick (25 mm), short (30 cm), and was dropped from chair height. The spectrogram indicates that bounces occurred at similar points in time to that of the wooden rod dropped from chair height (Figure 5.8). The overall sound level of the plastic rod in the spectrogram is lower than the wood and metal rods, and this is evident by the less

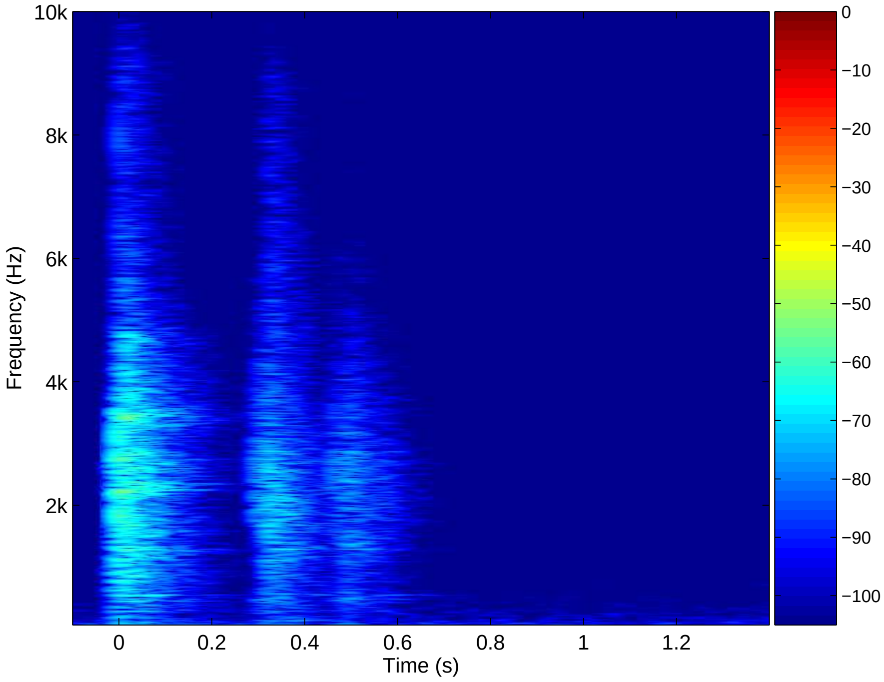


Figure 5.9: Spectrogram of a thick (25 mm diameter), short (30 cm), plastic rod dropped from chair height. Maximum levels are shown in red and minimum levels are shown in dark blue. The color decibel scale (relative) is matched to that of Figures 5.7 and 5.8.

intense colors. Although one can see faint horizontal lines and thereby potentially identify resonant frequencies, the strong resonances that were obvious for the metal rod and also clear for the wood rod are absent in this spectrogram.

5.3 Results

5.3.1 Group Accuracy

Mean subject response accuracies for the normal-hearing and hearing-impaired subjects, both in the unaided and aided conditions, are shown in Figure 5.10. All group results were above chance level. For material judgments, chance level was 33.3%, as there were three options from which the subjects could guess. For length and height judgments, chance level was 50%. Trials on which subjects marked the stimulus as inaudible (“I didn’t hear anything”) were counted as completely incorrect.

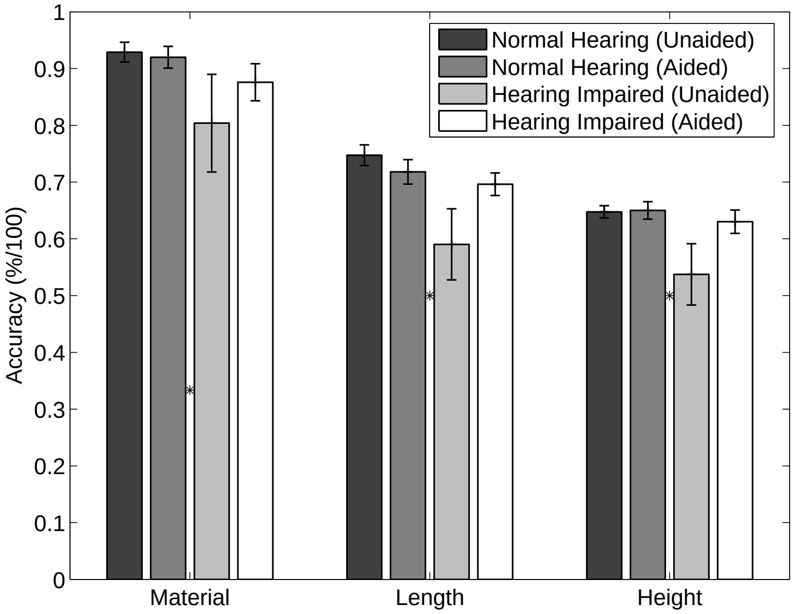


Figure 5.10: Comparison of accuracies for normal-hearing and hearing-impaired test subjects, each group with and without hearing aids, for each of the three properties that were judged (material, length, height). Standard errors, based on the individual subject mean accuracies, are shown at the top of each bar. Asterisks indicate chance level for each parameter.

For the statistical analysis of the data, the average accuracy of the four repetitions of each stimulus was used when performing counts of correct and incorrect answers. Instead of counting the number of correct responses as the actual total of correct responses per condition, the results were divided by four to account for the four repetitions. Because there were four times as many trials when considering the four repetitions and because these repetitions could not be considered to be independent of one another, a statistical analysis that did not take this into account may have falsely inflated differences and therefore an excessively high amount of power. By using an average of the results for the four repetitions, the repetitions are still able to provide useful data, but in an appropriately conservative manner.

A two-sided, corrected for continuity, Chi-square analysis (Siegel & Castellan, Jr., 1988) of the group data (total correct versus total incorrect for all subjects) indicated that the unaided normal-hearing subjects performed significantly better than the unaided hearing-impaired subjects for each of the three judged parameters (Material: $\chi^2(1) = 22.5$, $p < 0.001$; Length: $\chi^2(1) = 17.9$, $p < 0.001$; Height: $\chi^2(1) = 7.88$, $p = 0.005$). As a group, the hearing-impaired subjects performed worse than the normal-hearing subjects in all cases.

Next, a statistical comparison of unaided normal-hearing subjects to *aided* hearing-impaired subjects was made. The mean accuracy of this data is again shown in Figure 5.10. A two-sided Chi-square analysis of the group data indicated that the unaided normal-hearing subjects may have performed better than the aided hearing-impaired subjects for judgments of material (Material: $\chi^2(1) = 4.86$, $p = 0.03$), but not significantly better for length or height judgments (Length: $\chi^2(1) = 1.90$, $p = 0.17$; Height: $\chi^2(1) = 0.146$, $p = 0.70$). Therefore, only material judgments were found to be better, but without an extremely high degree of confidence.

A comparison of aided normal-hearing and aided hearing-impaired subjects was next made. As previously, a plot of the mean accuracy of this data is shown in Figure 5.10. Two-sided Chi-square analysis of the group data indicated that the aided normal-hearing subjects did not perform significantly better than the aided hearing-impaired subjects for judgments of material, length, or height (Material: $\chi^2(1) = 3.10$, $p = 0.079$; Length: $\chi^2(1) = 0.279$, $p = 0.60$; Height: $\chi^2(1) = 0.205$, $p = 0.65$). It appears that when wearing hearing aids, the two subject groups are most similar, whether due to the subjects being helped or hindered by the hearing aids.

Two-sided Wilcoxon signed ranks tests were conducted to determine whether the hearing aids used by the normal-hearing and hearing-impaired subjects resulted in a difference in performance for either of the individual subject groups. The Wilcoxon signed ranks test considers whether each subject of the chosen group performed better or worse in the aided condition, and from this result, weighted by the magnitude of performance change, a p value is generated which can be used to test the hypothesis of whether there is a performance difference between the conditions. For the normal-hearing subjects, there was no significant effect of the hearing aid (Material: $p = 0.29$; Length: $p = 0.17$; Height, $p = 1$). For the hearing-impaired subjects, if a level of 5% is used to determine significance, the hearing aid resulted in a borderline significant difference for length judgments ($p = 0.049$), a stronger difference for height judgments ($p = 0.008$), but did not change performance for material judgments ($p = 0.75$). Although the data is not normally distributed, paired t-tests were also conducted on the same data. For normal-hearing subjects, the results were similar to the signed ranks test results (Material: $p = 0.46$; Length: $p = 0.14$; Height: $p = 0.87$). For the hearing-impaired subjects, the paired t-tests also had similar results (Material: $p = 0.33$; Length: $p = 0.08$; Height: $p = 0.04$). Though the strict conditions for the paired t-test have not been met, the results serve to confirm the findings of the Wilcoxon signed ranks tests.

Parts of the analysis to this point have been strongly influenced by two hearing-impaired subjects with severe hearing losses, subjects LR and KC. Figure 5.10 is

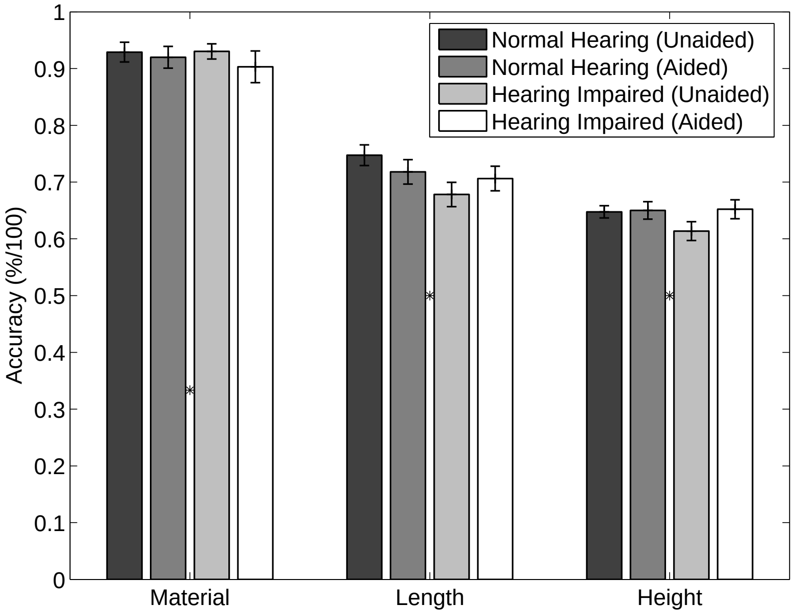


Figure 5.11: Comparison of accuracies for normal-hearing and hearing-impaired test subjects, each group with and without hearing aids, for each of the three properties that were judged (material, length, height). Data from severely-impaired subjects LR and KC have been excluded. Standard errors, based on the individual subject mean accuracies, are shown at the top of each bar. Asterisks indicate chance level for each parameter.

plotted again as Figure 5.11 with the data of hearing-impaired subjects LR and KC excluded from the analysis. The ten remaining hearing-impaired subjects were no worse than their unaided normal-hearing counterparts at the material, length, or height judgments, even when not wearing their hearing aids (Material: $\chi^2(1) = 0.011$, $p = 0.92$; Length: $\chi^2(1) = 3.17$, $p = 0.075$; Height: $\chi^2(1) = 0.597$, $p = 0.44$). Results were similar when comparing unaided normal-hearing listeners to aided hearing-impaired listeners (Material: $\chi^2(1) = 1.00$, $p = 0.32$; Length: $\chi^2(1) = 1.06$, $p = 0.30$; Height: $\chi^2(1) = 0.0008$, $p = 0.98$) and when comparing aided normal-hearing listeners to aided hearing-impaired listeners (Material: $\chi^2(1) = 0.333$, $p = 0.56$; Length: $\chi^2(1) = 0.0495$, $p = 0.82$; Height: $\chi^2(1) = 0.0014$, $p = 0.97$). For those unaided hearing-impaired subjects who were at least able to consistently hear the stimuli, this task was therefore apparently no more difficult than for the normal-hearing subjects.

Wilcoxon signed ranks tests conducted on the hearing-impaired subject data, this time excluding the severely hearing-impaired subjects LR and KC, showed that the hearing aids worn by the subjects resulted in no performance change for material judgments ($p = 0.45$), no change in length judgments ($p = 0.19$), but possibly a slight performance improvement for height judgments ($p = 0.031$). Paired t-tests of the same data arrived at similar conclusions (Material: $p = 0.28$; Length: $p = 0.20$; Height: $p = 0.028$). In analysis of the data excluding severely impaired subjects LR and KC, the results again indicate that categorization of drop height may have been improved through the use of hearing aids, even though unaided performance was not proven to be worse for unaided hearing-impaired subjects than for unaided normal-hearing subjects.

The statistical comparisons described in this section are summarized in Table 5.6. The differences between these two analyses are revealing. The results are rather different depending on whether the hearing-loss severity is capped. Unless otherwise stated, the data of the two severely impaired subjects is included in the remaining results.

Table 5.6: Summary of statistical comparisons made between normal-hearing and hearing-impaired listeners in unaided and aided conditions. Rightmost columns indicate the group and condition found to be statistically better for the comparisons indicated in the left column. Results are shown for the comparisons including all hearing-impaired subjects ("All Subjects") and for the comparisons excluding the subjects with severe impairments ("No Severe Losses"). Asterisks indicate degree of statistical significance. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. Comparisons in which no statistically significant difference was found are denoted by a hyphen ("–").

Group/Condition	vs. **Group/Condition**	**Better Group/Cond. (All Subjects)**	**Better Group/Cond. (No Severe Losses)**
Normal, Unaided (NH-U)	vs. Impaired, Unaided (HI-U)	Material: NH-U*** Length : NH-U*** Height : NH-U**	Material: – Length : – Height : –
Normal, Unaided (NH-U)	vs. Impaired, Aided (HI-A)	Material: NH-U* Length : – Height : –	Material: – Length : – Height : –
Impaired, Unaided (HI-U)	vs. Impaired, Aided (HI-A)	Material: – Length : HI-A* Height : HI-A**	Material: – Length : – Height : HI-A*
Normal, Unaided (NH-U)	vs. Normal, Aided (NH-A)	Material: – Length : – Height : –	Material: – Length : – Height : –
Normal, Aided (NH-A)	vs. Impaired, Aided (HI-A)	Material: – Length : – Height : –	Material: – Length : – Height : –

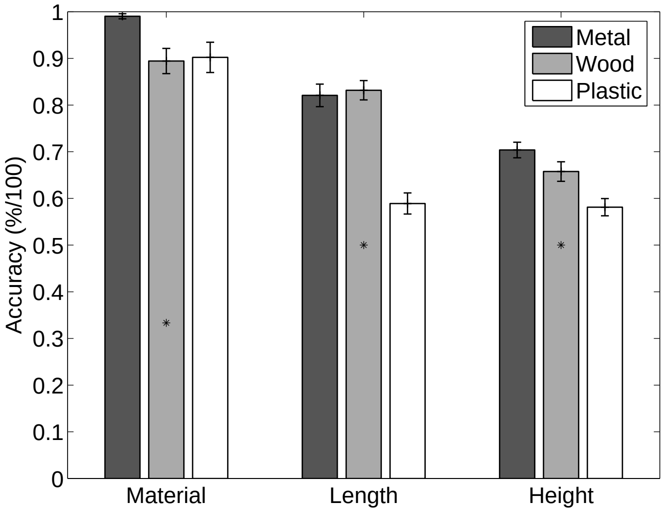


Figure 5.12: Average accuracies by material for unaided normal-hearing subjects. Standard errors, based on the individual subject mean accuracies, are shown at the top of each bar. Asterisks indicate chance level for each parameter.

Stimulus Specific

Parameter judgments for different materials were not of the same difficulty. Judgment accuracies for unaided normal-hearing subjects are shown in Figure 5.12 for each of the three materials. For example, it is apparent that judgments of length were more difficult for the plastic rods than the metal and wooden rods. Height judgment accuracies also suggested that listeners have more trouble judging the height from which plastic rods were dropped. Judgments of material for metal, wood, and plastic rods were very accurate. Metal was nearly always recognized correctly.

Parameter judgment accuracies for unaided hearing-impaired subjects are shown in Figure 5.13 for each of the three materials. As was the case for normal-hearing subjects, it is apparent that judgments of length were more difficult for the plastic rods than the metal and wood rods. Results for the aided hearing-impaired subjects reflected similar tendencies but with better performance. Note that length judgments of plastic rods made by unaided hearing-impaired subjects fall *below* chance level.

This can partly be explained by the fact that when the stimuli were inaudible to the subjects, the responses were counted completely incorrect. However, examination of the length estimates for short and long rods suggests a further explanation of why length judgments of plastic rods fell below chance level. Figure 5.14 shows

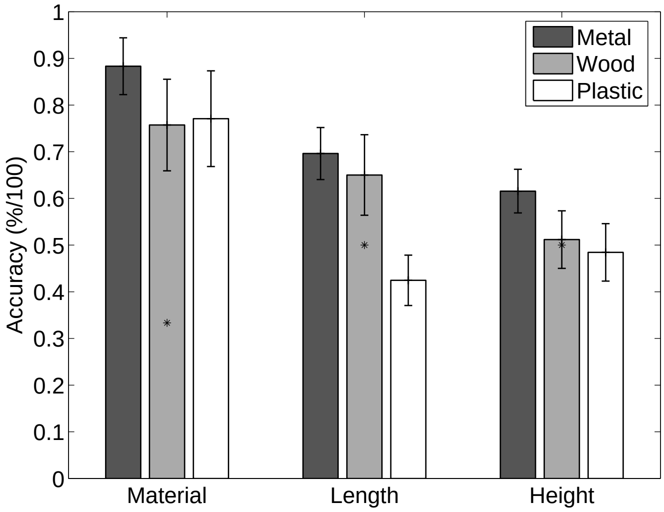


Figure 5.13: Average accuracies by material for unaided hearing-impaired subjects. Standard errors, based on the individual subject mean accuracies, are shown at the top of each bar. Asterisks indicate chance level for each parameter.

the average length-judgment accuracy, by material, for unaided hearing-impaired test subjects. It can be seen that for judgments of plastic rods, it is apparently much more probable that the subject was correct when presented with short rods than with long rods. This is likely because of a bias towards choosing “short”. It is hypothesized that this is because the sounds produced by dropping these rods were lower in sound pressure level than the other stimuli (see Table 5.4 on page 117). Based on comments made by test subjects, quiet sounds were likely to be interpreted as coming from sound events involving shorter rods. A bias was still present when the hearing aid users wore their hearing aids, though the accuracies of judgments made for both short and long stimuli were greater.

Bubble plots of confusion matrix data are able to provide a detailed picture about the confusions made by the test subjects. Figure 5.15 shows a bubble plot of the judgments made for all of the stimuli presented to hearing-impaired subjects in the unaided condition. The material, length, and height of the actual stimulus are compared to the response material, length, and height. The stimulus and response codes are formed from the first letter indicating material type (“M”=Metal; “W”=Wood; “P”=Plastic), the second two characters indicating the length of the rod in centimeters (“30”=Short (30 cm); “55”=Long (55 cm), and the last letter indicating from which height the rod was dropped (“S”=Shelf Height; “C”=Chair Height). As would be expected, short plastic rods dropped from chair height were the stimuli that were most often inaudible to the listeners. Particularly for plastic

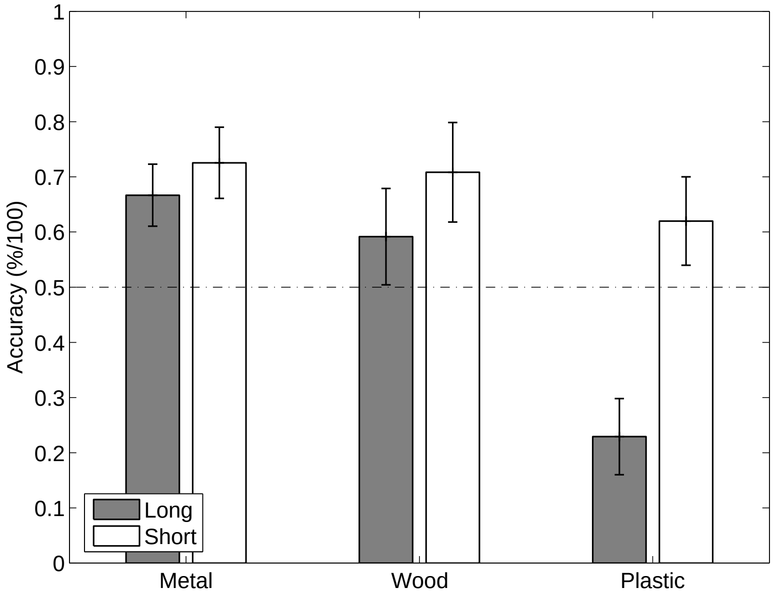


Figure 5.14: Short and long rod average length accuracy for unaided hearing-impaired subjects. Standard errors, based on the individual subject mean accuracies, are shown at the top of each bar. Chance level (50%) is shown by the dash-dotted line.

rods, one can observe that rods dropped from shelf height were often believed to be dropped from chair height, but not as frequently the other way around. From inspection of additional bubble plots (not shown), it was also found that narrow diameter (16 mm) rods were more frequently confused for being dropped from chair height than large diameter (25 mm) rods. It is hypothesized that this is because listeners may have based their drop height judgments with excessive weight on the intensity of the stimulus. The more massive, large diameter rods produced a more intense sound than the small diameter rods and therefore may have been interpreted as being dropped from higher heights.

Inspection of the confusion-matrix bubble plot in Figure 5.15 also reveals that for plastic rods of any length, a response of short ("30") and chair height ("C") was common. These confusions were less pronounced for wood and metal rods. For the plastic rods, subjects apparently had greater difficulty in utilizing cues for length and height perception. These cues may have been more subtle for plastic, but the subjects were also apparently biased in their responses.

As has been observed in other studies (Kunkler-Peck & Turvey, 2000; Tucker, 2003; Giordano & McAdams, 2006), most material confusions for normal-hearing listeners occurred between wood and plastic (Unaided: 6.6%). A very small number of

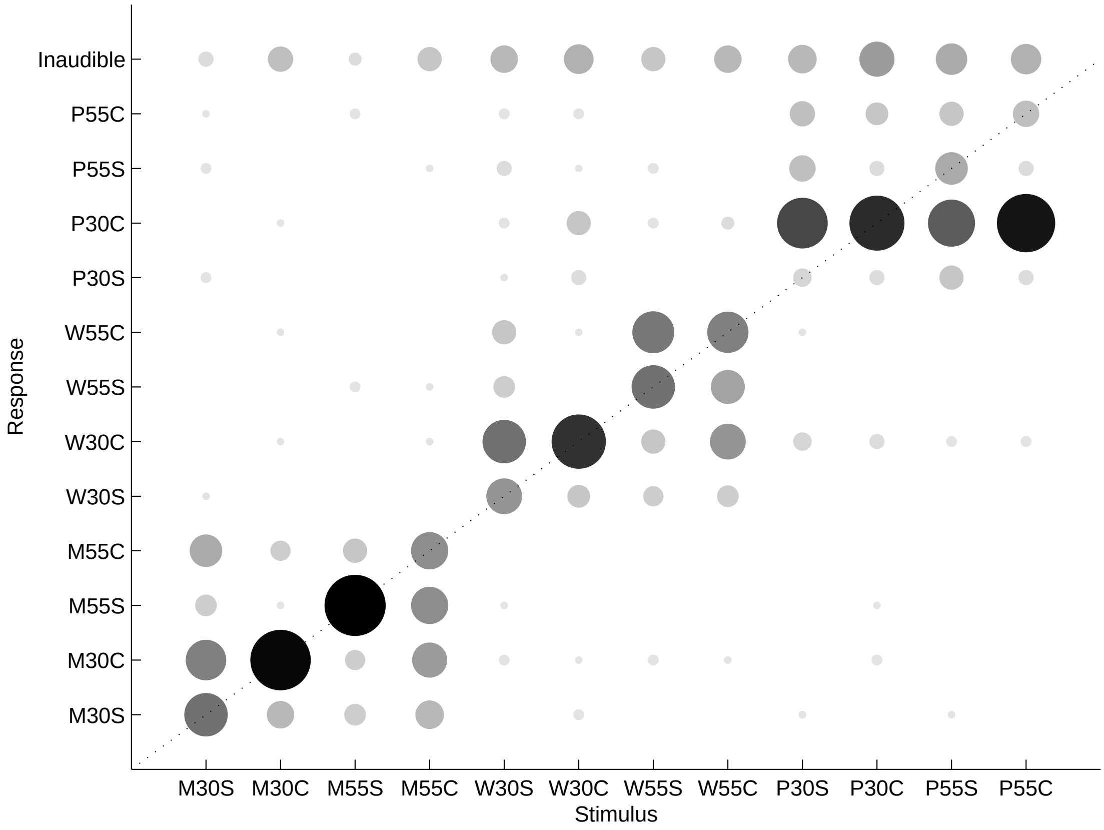


Figure 5.15: Bubble plot of confusions made by unaided hearing-impaired subjects between all stimuli, with thin and thick rods combined. Larger, dark bubbles represent a large number of responses, while few responses are represented by small light-colored bubbles. Bubbles falling on the dotted line are correct answers. Legend: “M”=Metal, “W”=Wood, “P”=Plastic ; “30”=Short (30 cm), “55”=Long (55 cm); “S”=Shelf Height, “C”=Chair Height.

confusions occurred between other materials, both when normal-hearing listeners wore hearing aids and when they did not (Unaided: 0.5%; Aided: 0.8%). Confusion combinations between materials other than wood and plastic also occurred with the hearing-impaired subjects, but again in relatively small amounts (Unaided: 2.6%; Aided: 0.4%).

Response Times

The median response times for each subject were examined. The median was used for calculations, as opposed to the mean, in order to reduce the influence of the (rare) trials on which the subjects stopped to ask questions. Median response times were 5.7 s for hearing-impaired subjects and 5.0 s for normal-hearing subjects. Normal-hearing subjects had response times ranging from 3.0 s to 7.3 s with a 1.3 s standard deviation of the mean. Hearing-impaired subject response times ranged from 2.5 s to 10.6 s with a 2.3 s standard deviation of the mean. Note that the fastest response times for hearing-impaired subjects were due to the stimuli being inaudible, and the subject quickly indicating this with a single mouse click.

The higher maximum response time with the hearing-impaired subjects can partly be explained by the mouse skill-level differences between the subjects. Some of the older, hearing-impaired subjects had rarely used a mouse before. Because of this, a check of differences between the two groups' response times was not conducted. However, t-tests comparing the mean response time for the individual subject groups between the aided and unaided conditions showed that hearing-impaired subjects performed moderately slower when performing the test with their hearing aids ($p = 0.044$). This result should be interpreted with consideration for the fact that inaudible stimuli, of which there were more in the unaided condition, generally allowed for faster response times. When the data of the two severely hearing-impaired subjects (those subjects who accounted for most of the inaudible responses) was excluded from the analysis, response times were no different for the hearing-impaired subjects when wearing their hearing aids than when not wearing their hearing aids ($p = 0.15$). Response times from the normal-hearing subjects showed no effect from the use of the hearing aids ($p = 0.60$).

A correlation coefficient was calculated in order to determine whether performance accuracy was related to response time. For this purpose, an average accuracy of all three parameter judgments (material, length, and height) was calculated for each subject and for each condition. The entire set of resulting accuracies was compared to the set of response times for those accuracies. Unaided conditions of hearing-impaired subjects LR and KC were excluded because of their unnaturally fast response times resulting from the fact that so many trials were quickly marked inaudible by way of a single mouse click. The correlation coefficient was -0.29, indicating little linear correlation between performance accuracy and response time. The spread in the data was large.

Training and Order Effects

For both the normal-hearing subjects and the hearing-impaired subjects, Wilcoxon signed ranks tests showed that there was not a significant change between the overall accuracy of the subjects between the first and last sessions (Normal Hearing: $p = 0.13$; Hearing Impaired: $p = 0.35$). Paired t-tests showed similar results (Normal Hearing: $p = 0.31$; Hearing Impaired: $p = 0.89$). Therefore, there was apparently no significant training effect that resulted in an improvement or worsening in performance between the first and last sessions.

There was also no evidence that the order in which test subjects did the test conditions (either the aided sessions first or the unaided sessions first) made a significant difference – neither for the normal-hearing subjects nor the hearing-impaired subjects. Performance of normal-hearing subjects who completed the unaided conditions first was not different from the performance of those subjects who completed the unaided condition last ($p = 0.64$), nor was the performance of normal-hearing subjects who completed the aided condition first different from the performance of the subjects who completed the aided condition last ($p = 0.58$). Similarly, hearing-impaired subject performance was no different if unaided in the first half versus the second half of the test ($p = 0.44$), nor if aided in the first half versus the second half ($p = 0.16$). Paired t-tests showed similar results.

5.3.2 Individual Subject Accuracy

The above analyses have examined group performance, but examination of the individual results shows some interesting effects. Although group results were significantly above chance level in most cases, individual results were not always above chance level. The two severely impaired test subjects (LR and KC) even performed significantly worse than chance level for all three questions in the unaided condition, due to the fact that many of the stimuli were marked inaudible, and these trials were counted completely incorrect. As an example, Figure 5.16 shows the length categorization accuracies for each of the unaided hearing-impaired subjects. Figure 5.17 shows the length categorization accuracies for the same subjects in the *aided* condition.

Length estimates for the two subjects were greatly improved to above chance level

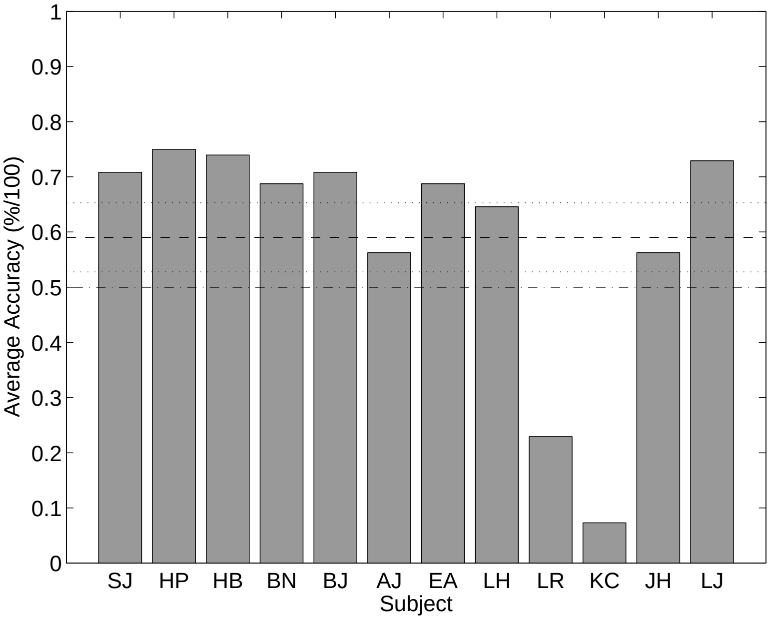


Figure 5.16: Average length accuracy for unaided hearing-impaired subjects. Chance level (50%) is shown as the dash-dotted line. Overall mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

(LR: $\chi^2(1) = 13.5$, $p < 0.001$; KC: $\chi^2(1) = 5.1489$, $p = 0.023$), as seen when comparing Figure 5.16 (unaided) to Figure 5.17 (aided). With their hearing aids, subjects LR and KC also performed above chance level for material judgments (LR: $\chi^2(1) = 6$, $p = 0.014$; KC: $\chi^2(1) = 52.128$, $p < 0.001$). However, subjects LR and KC still had trouble with height estimates. Neither of their results were different from chance level (LR: $\chi^2(1) = 1.042$, $p = 0.31$, KC: $\chi^2(1) = 0$, $p = 1$). Plots illustrating material and height estimates for each subject will not be shown here but may be found in Appendix C. By making the sounds audible to these two subjects, the hearing aids had a great impact. Considering for a moment that even if the accuracy of the parameter judgment is ignored, these subjects were at least helped by their hearing aids to become aware of sound events, a fact which in and of itself is valuable.

All other hearing-impaired subjects were above chance level in their estimates of material in both the unaided and aided conditions ($p < 0.05$). Not all hearing-impaired subjects were above chance level in their length and height estimates for a given condition, but no obvious tendencies were observed for the individual subjects to necessarily do better or worse with their hearing aids. Differences could have also been caused by training effects in individual test subjects, therefore it

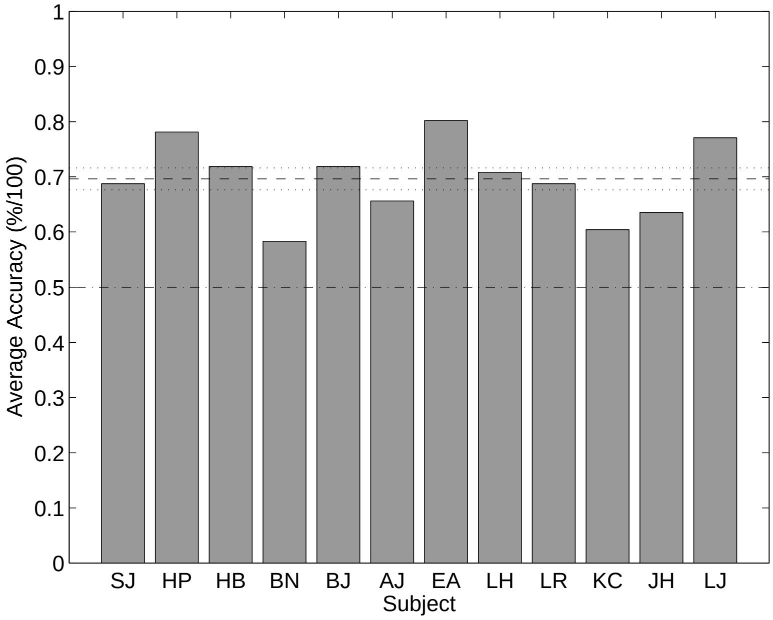


Figure 5.17: Average length accuracy for aided hearing-impaired subjects. Chance level (50%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

was deemed inappropriate to draw conclusions for individual subjects concerning whether the hearing aids were responsible for changes between conditions. Only in the previously mentioned cases of the hearing aids clearly helping the severely impaired subjects were the results clear on an individual basis.

For material judgments, all normal-hearing subjects performed above chance level in both the aided and unaided conditions. Unaided length judgment accuracies were also completely above chance for the normal-hearing subjects, but one subject (GP) scored at chance level in the aided condition. The aided condition was done as the second half of the test for this subject, so it is unlikely that training effects could explain the result. After the listening test, the subject mentioned that he changed listening strategies during the test, and this may have affected the results of the aided session. The subject did not elaborate on the strategies used. However, inspection of the subject's response times (not shown) hint that he spent a greater than usual amount of time considering his responses. Length judgment accuracies for this subject and the other unaided normal-hearing subjects are shown in Figure 5.18.

Height judgment accuracies for the normal-hearing subjects were examined next. All subjects scored above chance level (50%) in at least one condition ($p < 0.05$).

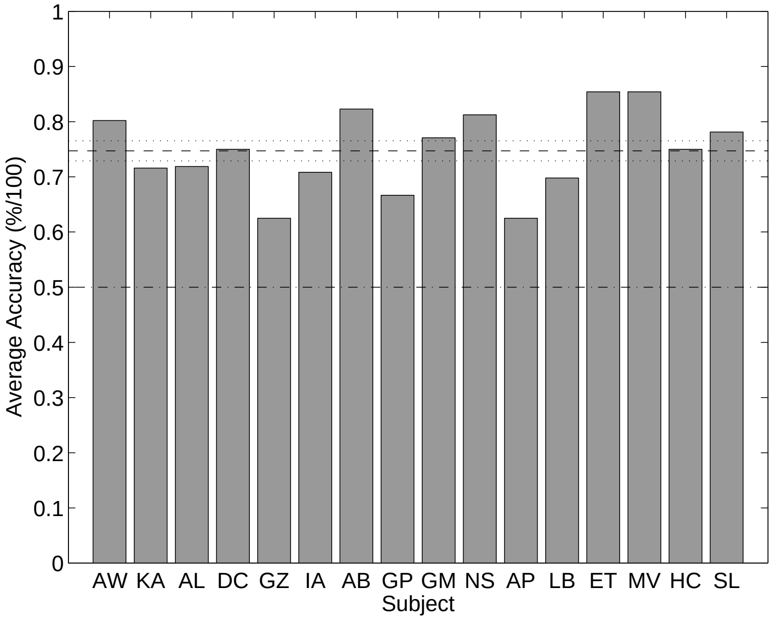


Figure 5.18: Average length accuracy for unaided normal-hearing subjects. Chance level (50%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

However, some subjects scored at a level that could not be said to be different from chance level in one of the conditions of their tests (Unaided: KA, GZ; Aided: DC, NS, HC). Performance was not necessarily improved for the condition which the subject completed second, but the data was not profound enough for drawing strong conclusions concerning whether the hearing aids were responsible for helping or hurting the performance of these particular subjects on an individual subject basis.

It should be noted that performance of individual subjects was not sharply divided by classification as “hearing impaired” or “normal hearing”. There were subjects in the hearing-impaired group who performed better than some of the normal-hearing subjects performed, and thereby normal-hearing subjects who performed worse than some of the hearing-impaired subjects performed, even when the hearing-impaired subjects were unaided. Visual comparison of Figures 5.16 (unaided hearing-impaired subjects) and 5.18 (unaided normal-hearing subjects) demonstrates this for length judgments. The same effect could be seen for material and height judgments (see Appendix C).

5.4 Discussion

Although it was not the primary goal of the experiment, the finding that the normal-hearing subjects were very good at being able to discern between metal, wood, and plastic confirms the results of Kunkler-Peck & Turvey (2000). As was done in the experiments conducted here, Kunkler-Peck & Turvey presented their stimuli live. Their stimuli were plates of different materials being struck. However, the findings of the present experiment are in opposition to those of Giordano & McAdams (2006) who found that normal-hearing test subjects had trouble hearing the difference between wood and plastic. This problem was not observed in the experiments described in this Chapter. Even the full group of hearing-impaired subjects, including those with severe losses, were found to make judgments of the material of wooden and plastic rods that were each accurate more than 75% of the time. Unaided normal-hearing subjects rarely confused wood and plastic.

In their experiments, Giordano & McAdams used monophonically recorded stimuli that were presented to subjects over headphones. They suggested that the differences in presentation method may account for the differences between their results and those of Kunkler-Peck & Turvey (2000). There is therefore mounting evidence, further supported by the experiments described in this chapter, that subjects are better at perceiving material when listening to authentic, live stimuli. Monophonic recordings appear to prevent accurate estimates of the abilities of listeners to hear the materials of which objects are made. This suggestion supports the findings presented in Chapter 4 that monophonic recordings are poorly able to represent real sound events. In addition to perceptually relevant information on size, apparently information on material may be also compromised by monophonic recordings. It is not immediately clear why material would be poorly represented by a monophonic recording. Further investigation is needed to verify this observation and to explain why it occurs.

The results of this experiment additionally confirm those of Tucker (2003), who found that listeners were better capable of estimating the sizes of metal objects than then sizes of plastic objects. Both the normal-hearing and the hearing-impaired subjects showed similar difficulties in perceiving the length of plastic rods. The investigation of Tucker involved asking subjects about the size of struck plates, and so it appears that material perception may be similar for both struck plates and dropped rods – two types of impact sounds.

As a group, it was shown that unaided hearing-impaired test subjects are worse than normal-hearing subjects at discerning material, length, and the height from which a rod is dropped. However, large differences between individual subjects were found. Data analysis excluding two severely impaired subjects showed that the remaining ten unaided subjects were not worse than normal-hearing subjects performing the same task.

For the subject group as a whole, the hearing aids worn by the hearing-impaired subjects did not completely restore to normal the group's abilities to discern characteristics of impact sounds. Material judgments were found to be slightly worse for aided hearing-impaired subjects than for unaided normal-hearing subjects. Their hearing aids were not shown to be of benefit for material categorizations. The hearing aid processing was apparently insufficient or inappropriate for aiding the group in making judgments about material. On the other hand, aided hearing-impaired subjects' estimates of length and the height from which the rods were dropped were not found to be statistically different from estimates made by unaided normal-hearing subjects. The hearing aids were of significant benefit to the hearing-impaired subjects, as a group, for making height categorizations and of slightly significant benefit for making length judgments.

For material perception, it is evident that some parts of the hearing losses may not have been fully restored. The average hearing-impaired listener equipped with hearing aids was somewhat less capable of getting information from the stimuli than was the average unaided normal-hearing person. The difference of approximately 5% in aided hearing-impaired subject accuracy versus unaided normal-hearing subject accuracy represents the average result of all subjects of the respective groups. It should be emphasized that some of the hearing-impaired subjects may have experienced a large benefit from their hearing aids, while others may have had no benefit or even experienced a decrease in performance when wearing their hearing aids, but sufficient data for such individual comparisons was not available. Only for the results of the two severely impaired subjects could specific conclusions be drawn with adequate certainty.

The fact that unaided normal-hearing listeners were on average slightly more successful than the full group of aided hearing-impaired subjects at making material judgments could be because the hearing aids did not do a sufficient job of presenting the relevant information for the recognition of material. This of course may not be simply a matter of the hearing aid not making the information audible to the subjects – that is, of reasonable intensity – but that the signal may have been compromised by the hearing aid. For example, the dynamics of a hearing aid compression system (e.g., its attack and release time constants) may distort an impact sound in a way that is detrimental to perception. Processing that is more sophisticated than basic amplification may be required to convey "material", as information, to a person suffering from hearing impairment. This point may apply, in particular, to listeners suffering from auditory processing disabilities other than simply raised thresholds (e.g., decreased frequency resolution or decreased temporal resolution). These issues appear to be primarily relevant to those test subjects with severe hearing impairments, as the performance difference for material judgments between unaided normal-hearing subjects and aided hearing-impaired subjects disappeared when the two hearing-impaired subjects with the most severe hearing impairments were excluded from the analysis.

Performance of *aided* hearing-impaired subjects compared to *aided* normal-hearing subjects indicated that the normal-hearing subjects were not better at perceiving material, length, or height while wearing hearing aids. The difference in results when comparing normal-hearing listeners in the unaided and aided conditions with the aided hearing-impaired listeners, may suggest that normal-hearing listeners were worse with hearing aids, but this could not be shown statistically. Aided normal-hearing subject performance was neither better nor worse than unaided normal-hearing subject performance. Either the critical information for success at the task was not aided or compromised by the hearing aid, or the signal may have been degraded, but the normal-hearing listeners were able to deal with the degraded signal in a manner sufficient to extract the necessary information.

Large performance differences were seen not only among the hearing-impaired subjects, but also for the normal-hearing subjects. The labels of “normal hearing” and “hearing impaired”, used to describe subjects with better or worse hearing, are gross categorizations of populations that may actually vary widely in their hearing abilities. No psychophysical or classical audiometric tests other than pure-tone threshold measurements were done to assess the functional auditory abilities of the test subjects in this test. Although it is very common, pure-tone audiometry is known to describe only a small part of the functionality of the ear as a sensory organ. For example, one subject used in this test with a severe high-frequency hearing loss has been found in the research of Papakonstantinou (2005) to have pure tone frequency discrimination abilities on par with normal-hearing listeners (Zeng, Kong, Michalewski & Starr, 2005). With a large group of subjects and sufficient time, additional tests may help explain performance differences between subjects. However, such an investigation may be unwarranted at this time, while only exploratory work is being done.

Because the test subjects used in this experiment were different not only in their hearing abilities, but also in their ages, differences could also be due to the fact that the normal-hearing subjects possessed greater cognitive abilities to recognize material. The hearing-impaired subjects were all older than the normal-hearing subjects were. They may have had cognitive deficiencies resulting from old age that also influenced their performance in the task. Cognitive problems may have also prevented some of the hearing-impaired subjects from receiving full benefit from their hearing aids (Lunner, 2003; Gatehouse et al., 2003). On the other hand, the greater age of the hearing-impaired subjects could be an asset to them – they are likely to have more experience listening to impact sounds, simply by having lived for a longer period than the young subjects have. These issues were not controlled, and it is unclear if there would have been any benefit in doing so. Poulsen & Keidser (1991) found in a test of normal-hearing young people compared to normal-hearing elderly people that there were no differences in speech recognition performance in noise. Though it is unclear if those results could apply to the present investigation, apparently any cognitive differences that may have been present were not a problem in that particular task. Incidentally, the two most severely hearing-impaired subjects tested in this experiment were young compared to rest of the hearing-impaired subjects: 55 and 57 years old.

In their daily lives, all hearing-impaired subjects described in this chapter wore their hearing aids for at least four hours per day, and they had done so for a minimum of a few years. Although this check was made to insure that the subjects included in the test were experienced hearing aid users, it says nothing about how active they are in their lives or of the quantity and types of sounds to which they are exposed. For instance, a sedentary lifestyle may not give the hearing-impaired subjects enough experience with their hearing aids to take advantage of them. However, their eagerness to participate in the listening test is at least a small indicator of their activity level.

Testing with stimuli dropped behind the test subjects most likely made the task more difficult than if the rods had been dropped in front of the subjects. Some of the hearing aids used by the hearing-impaired subjects may have had directional microphones or processing designed to quiet sounds coming from behind the user, and these subjects in particular would have likely suffered from having stimuli presented behind them. However, it was chosen to drop the rods behind the test subjects because, in addition to practical reasons, this location is in fact a typical one in which hearing is of great importance – a position that is out of sight. It should be emphasized that many, if not most, environmental sounds do not occur directly in front of the listener experiencing them.

From a survival point of view, an idea to make the judgment of drop height more ecologically relevant could be to drop rods from above head height and below head height. Rods falling from above head height could be of greater concern to subjects than those simply falling from a height below their head. This was not done in this experiment due to the already-high maximum sound levels with some of the stimuli, but could be done if different rods were chosen to be tested.

Although no subject reported it when asked, “unwanted” noises such as the rods radiating sound as they rolled off the dropping apparatus could have given listeners cues to judge the attributes, which were inquired about, that did not result from the impact itself. This was undesired, but there is no reason to suspect such cues may not also be available in an everyday listening situation. For making group comparisons and considering the fact that these sounds were not consciously detected by test subjects, this issue was deemed of little importance.

The response system used in this test appeared to be easy for the subjects to use, but it was not particularly intuitive from an ecological perspective. Unlike in the experiment described in Chapter 4, where the subjects were physically involved in their response, providing an answer that physically represented the stimulus that they were asked to physically describe, the computer-interface with which the subjects responded in this experiment was not directly, physically related to the stimulus being described. Alternatives were considered, but it was decided that three separate response apparatuses, although potentially more intuitive in and of themselves, would be too cumbersome and confusing to operate in response on every trial.

In order to make the computer response system and types of questioning as simple and straightforward as possible, multiple-choice questions were asked, which each had a very small set of possible answers. Because of this, the high value of the chance level, particularly for judgments of length and height, may have limited the abilities to find differences between subjects. Greater differences may have been seen if the chance level accuracies, the levels at which guessing would result, had been lower.

The order of the trials and how the parameters varied between them may have affected the responses. Although no statistical investigation was performed concerning the matter, the sequence of the stimuli for a particular subject may have influenced the results. This influence is likely unavoidable in a test setting, but the Latin square based system of randomization has hopefully helped to spread the order effects, thereby making the group analysis still valid.

It was potentially a disadvantage of the test that hollow rods were not available in all materials, solid rods were not available in all materials, and further material properties could not have been controlled more stringently. As a hypothetical example, if subjects could hear that a rod was hollow, they may have used this cue as an indicator of material if it was also assumed by the subject that hollow wooden rods may not be as common as hollow plastic rods (e.g., pipes). Therefore, it is possible that material judgments were influenced by the perception of hollowness/solidity. This limitation should be kept in mind if one wishes to generalize the findings of this experiment.

For both ecological and practical reasons, stimuli were presented live in this test.

Based on the results of the experiment described in Chapter 4, the binaural recording and playback method was the only alternative presentation method considered. However, because it was desired that the hearing-impaired subjects use their own hearing aids during the test, and because headphone playback to ears already equipped with hearing aids is troublesome, live presentation was chosen. Employing the binaural technique without the use of headphones would have required either a special presentation technique requiring an anechoic chamber with the subject locked into a fixed position or consideration for each of the direct audio playback systems possible with the hearing aids brought in off the street by the hearing-impaired subjects. An alternative solution may have been a high-resolution surround-sound system like Ambisonics (e.g., Fellgett, 1974; Guastavino et al., 2005), but equipment for using such a technique was not available for this project.

The choice of live presentation constrained the types of questions asked of the test subjects in this test. It would have been more difficult, for example, to present live stimuli in which the surface on which the rods were dropped was varied, to change whether there was an obstacle between the stimulus and subject, or to vary the distance between the stimulus and test subject without giving away clues between trials. The three stimulus parameters that were asked about, length, material, and height of drop, were therefore chosen not only for their ecological significance, but for practical reasons as well.

There is no more realistic way to present stimuli than live. Just like in the real world, a subject may never hear strictly identical sound events twice. In an experimental situation using live stimuli, there may be slight differences between repeated presentations of the same stimulus. From an ecological perspective, this could be considered as an advantage. While there could be added variance in the data caused by changes in the stimuli between trials, the results should be most realistic. If miniscule differences between the performances of subject groups are of interest, then the presentation of recordings could be useful. Because only practically significant differences were of interest for this investigation, live presentation was a suitable choice. Exposing the subjects to multiple “angles” of the stimuli may even have had the effect of naturally improving performance. It was shown in the experiment described in Chapter 4 that, if anything, subjects performed better when the variation of live presentation existed between tests.

If the primary goal of the research experiment had been to find correlations between acoustical cues and perceived event properties, it may have been important to consider using recorded stimuli. Alternatively, binaural recordings with a nearby acoustic manikin could have been made during each subject’s listening test at the same time that the stimuli were presented live. This approach would maintain the benefits of live presentation while also allowing a later acoustical analysis on signals very similar to those heard by the subject.

More advanced statistical analysis techniques such as the use of generalized mixed effects models could possibly have been used to analyze the data collected in this experiment. However, considering the goals and size of the investigation, such methods were judged inappropriate and superfluous. In keeping with the philosophy that only practically significant differences were of interest, fairly conservative statistical tests have been used in the analysis. More advanced statistical methods may be useful for exploring the data for different purposes.

Using hearing-impaired test subjects has helped explain not only which deficiencies exist for the hearing impaired, but may also allow one to form ideas concerning which properties of the objects involved in an impact, and which properties of the impact itself, are less robustly transmitted. Such properties may be compromised in difficult listening situations, whether they are difficult due to a hearing impairment or for other reasons. The decision to use hearing losses of widely varying severity was advantageous in the sense that data from many different circumstances has been gathered. The differences between subjects have naturally varied the difficulty of the test, and results were collected at each of those difficulty levels. The findings are therefore general, but one could learn more about specific types of hearing losses by selecting subjects with only one type of loss, for example severe hearing impairments.

The large influence of two severely hearing-impaired subjects was enough to produce different results for the entire group. Because the results of the subjects with milder hearing losses could not be seen as clearly when the subjects with severe losses are included in the data, the analyses of both groups was useful. Observation of the differences illustrates that the severely impaired have big problems without their hearing aids, and that the hearing aid has an important place in their lives when it comes to the perception of everyday sounds. However, the result also illustrates that their hearing has not been restored to normal. Further tests using subjects with severe hearing losses may be fruitful if it is desired to improve the benefit of hearing aids. The presence of the data from these subjects helps to illustrate the “soft spots”, where errors may be made by the hearing impaired.

In this experiment, most errors were made by the severely hearing impaired, but not all impact sounds in the real world are of the same intensity or occur under the same conditions as those presented in this experiment. Therefore, it seems reasonable to assume that more than just the severely hearing impaired may have difficulties recognizing properties of many impact sounds. For this experiment, the sound levels used have simply set a threshold for “acceptable” hearing losses, but in everyday life, a much wider variety of impact sounds will be encountered. The choice of an appropriate sub-population of hearing-impaired subjects for further studies would therefore depend on the level of the sounds and likely on the amount and type of background noise.

Although it is unknown exactly what processing was done by the hearing aids of the two severely impaired subjects, comparison of the results with and without these subjects hints that providing audibility is an important start to transmitting everyday sound information efficiently to those with hearing impairments. Making an acoustic input sufficiently audible may not be the only important task of a hearing aid, but for these subjects it was likely an important aspect in detecting properties of the sound events. Insuring audibility is vital before any decisions can be made concerning the objects involved in interactions. Audibility of the initial impact as well as trailing bounces of dropped rods can assist in estimates of height as well as size. Similarly, the variation in the times required for the resonance of different materials to decay may be helpful in discerning material. It could therefore be important that the entire signal is audible, not just the initial impact.

If one wishes to study how to improve the perception of properties such as material, it would be reasonable to perform investigations using only those subjects who have trouble recognizing those properties. This of course will not give an accurate representation of the performance of the entire hearing-impaired population, but it would likely make finding solutions easier for the problem of improving the hearing of those with hearing impairments. The “noise” added by subjects in the hearing-impaired subject group, who for one reason or another do *not* have difficulty hearing the relevant property, could be reduced in this way, allowing researchers to more easily see the problems needing solutions and the benefits of potential solutions.

Chapter 6

General Discussion

> *We did a few more experiments, and I discovered that while bloodhounds are indeed quite capable [of using their sense of smell], humans are not as incapable as they think they are: it's just that they carry their nose so high off the ground.*
>
> -Richard P. Feynman (1984)

6.1 Summary of Key Results

The human auditory system involves more than the ear. An attempt to understand human hearing must account for this. The purpose and use of the auditory system must be considered from the beginning. An information-based approach to auditory perception, as has been used in the experiments described in this thesis, helps to provide a realistic picture of the use of the human auditory system.

The experimental results presented in Chapter 4 have confirmed those first reported by Carello et al. (1998) that listeners are fairly good at being able to hear the lengths of dowels dropped onto a hard surface. When performing the test while listening to stimuli generated live, test subjects were able to estimate the lengths of the wooden rods of lengths 15 to 120 cm, with an error magnitude of approximately 15%. The results of the experiment further suggest that the most popular presentation method for auditory experiments, namely the presentation of monophonic recordings via headphones, may be a poor choice for achieving accurate perceptual representations of sound events. When normal-hearing test subjects were asked to estimate the lengths of the eight wooden dowels from the sounds of the dowels being dropped onto a linoleum floor, monophonic recordings presented to test subjects over headphones were shown to result in larger length judgment errors than did live presentation. Size estimation errors made with monophonic recordings were more than 40% greater than errors made when listening to live stimuli, yielding an average estimation-error magnitude of approximately 21%. The results suggest that spatial auditory cues, so often neglected in hearing research, may provide information to listeners in this task. Furthermore, it is apparent that there are also other cues contributing to length perception. This is made clear by the fact that even with only monophonic recordings, listeners were able to produce reasonable estimates of the lengths of the rods. Spatial cues are likely but a piece of the puzzle.

These findings imply that the judgment accuracies reported in many previous auditory experiments, in which monophonically recorded or synthesized stimuli have been used, may be underestimates of the true abilities of humans. Physical size perception, be it of large or small objects, is apparently compromised for subjects listening to monophonic recordings. These findings may be most acutely relevant for tasks in which perception of size, object motion, object facing angle, or relative distance could possibly contribute to perception. One then may wonder for which types of stimuli these cues could be important. The philosophy that everyday sounds are extremely rich in information would suggest that these cues could be relevant to many types of sound events.

In addition to the accuracy of the information perceived, the fact that a spatially accurate presentation method was demonstrated to be capable of producing subjectively realistic stimuli may have further ramifications. A test subject who is unaware that he or she is listening to a reproduction is likely further capable of focusing on the sound event and not on the reproduction method used in the test environment, improving the ecological validity of the test. For these reasons, the results of the experiment presented in Chapter 4 indicate that live presentation and the headphone presentation of properly equalized binaurally recorded stimuli are reasonable options for presenting perceptually realistic stimuli to normal-hearing test subjects participating in experiments.

Live presentation and an additional sound-processing device, the hearing aid, were used in further experiments described in Chapter 5. The abilities of hearing-impaired subjects to perceive properties of impact sound events were compared to the abilities of normal-hearing subjects in a set of experiments in which rods were dropped on the floor of the same test room used for the experiments presented in Chapter 4.

It was hypothesized that hearing-impaired subjects would have a more difficult time than normal-hearing subjects at extracting information from the acoustic signals resulting from impact sounds. The experiment described in Chapter 5 suggested that an important part of this difficulty may simply be due to a lack of sufficient audibility. An analysis was done for two groups of hearing-impaired subjects: 1) all subjects, including two with severe hearing losses and 2) only the subjects from the complete group that had mild to moderate hearing losses (i.e., two subjects with severe losses were excluded from the second group). When hearing-impaired subjects with mild to moderate hearing losses listened to rods being dropped onto a floor, their abilities to judge the lengths of the rods, the materials from which the rods were made, and the heights from which the rods were dropped onto the floor, were no worse than normal-hearing subjects, even when the hearing-impaired subjects were not wearing their hearing aids.

While the group of hearing-impaired subjects with mild to moderate losses were just as good as the normal-hearing subjects, the same was not true for the complete group of hearing-impaired subjects, including the severely hearing impaired. The unaided group, as a whole, were worse off than the normal-hearing subjects at perceiving the materials of the rods, the lengths of the rods, and heights from which the rods were dropped onto the floor. Even with their hearing aids, the entire group of hearing-impaired subjects remained worse at making material judgments. This suggested that research attempts intending to provide information useful for improving hearing aids should focus on severely hearing-impaired subjects, or subjects for whom an alternative task is found to be difficult.

Until now, questions like those asked in this experiment have rarely been asked of hearing-impaired listeners. Questions of this type, which are concerned with information provided by everyday sounds, have commonly been ignored in hearing research. Speech is the primary form of sound as information that has classically been considered. However, the experiment described in Chapter 5 has produced basic knowledge of how well a group of hearing-impaired listeners can perceive material, length, and height. The results have also demonstrated that while hearing aids are beneficial in some respects related to impact sound perception, they do not provide improvement in all cases.

6.2 Hearing Impairment and Information-Based Perception

Of importance to correcting hearing impairments is addressing the possibility that hearing-impaired people are less capable than normal-hearing people at acquiring information from sounds in their environments. Sensations are of only indirect interest, if of any interest. As suggested in Chapter 3, an attempt to help those with hearing impairments should include constant consideration for this. In Section 3.3, an information-based approach to studying hearing impairment was suggested. An outline of steps to help the hearing impaired was described on page 75. This thesis has focused on the first two steps of the outlined procedure, but has produced results relevant for many of the other steps as well.

To briefly recap the steps:

1. Determine if the hearing impaired are worse at perceiving information.
2. Identify ecologically significant problems of the hearing impaired.
3. Investigate how the normal hearing acquire this information.
4. Find ways in which the impaired auditory system is different.
5. Assess why the hearing impaired have difficulties acquiring information.
6. Consider how information transmission may be restored.

Step one proposed that it would be useful to first determine if there is a difference between the amount or accuracy of information that a hearing-impaired person gets from his or her environment compared to a normal-hearing person. The results from the experiment presented in Chapter 5 address this point and demonstrate that there are differences between groups of subjects, but that these differences are sensitive to precisely how the subject groups are defined.

Step two suggested that if the hearing impaired have been found to be incapable of obtaining the same information from their environments as normal-hearing people obtain, then this information should be identified and consideration should be given to its ecological relevance. Assuming one started with ecologically motivated questions in step one, the ecological relevance of the information should be somewhat already secured. However, it is then necessary to assess the ecological relevance of the magnitude of the reduction in information received. Focus should be placed on areas with the largest, ecologically relevant performance differences between the hearing-impaired listeners and normal-hearing listeners. Alternatively, it may be necessary to focus on those differences for which solutions are possible, but this may only be known after further research. Information reception differences found in Chapter 5 for unaided normal-hearing and hearing-impaired listeners suggest that there were hearing-impaired subjects who had ecologically significant difficulties in hearing the materials of rods, the lengths of rods, and the heights from which the rods were dropped onto the floor. However, not all of the subjects had difficulties. Further research should consider using subjects who have hearing losses similar to the losses of the subjects who performed poorly in this experiment.

It should be verified that the results of the tests concerning material, length, and height perception can be generalized and are not just specific to the tests conducted for this project. If the results are general, step three should be undertaken. Step three suggests that it would then be helpful to understand how normal-hearing listeners perceive the information that was selected in step two for further study. By following this series of steps, the experiments producing results that describe *what* people hear can be used to guide research concerning *how* people hear. Without first understanding for what purpose people use their auditory systems, studies of how people hear may have no ecologically relevant direction. Ecological considerations can and should provide that direction. Once the direction has been clarified by steps one and two, investigations can begin concerning how the normal auditory system extracts information. With a little luck, some of the results of classical psychoacoustics may become relevant at this point, but it is likely that many more ecologically motivated investigations are needed.

Chapter 4 described experiments in the perception of length by normal-hearing subjects that may provide information helpful for step three. The results of the experiment suggested that spatial cues may be one type among many important types of cues for the perception of size, but this hypothesis should be confirmed. The historic and future research of others is needed in understanding height and material perception, and in further understanding length perception. For example, as described on page 24 of Chapter 2, Warren, Jr. et al. (1987) found evidence that listeners were sensitive to temporal patterning that could be informative about height for bouncing objects. It is likely that there are other cues that also contribute to information concerning height, but their results may provide a starting point for considering height perception in hearing-impaired listeners. Other researchers have made hypotheses concerning how normal-hearing people hear the nature of a material, but the explanations to date are not complete. Most recently, Giordano & McAdams (2006) stated that many acoustic signal properties such as signal duration, damping, frequency content, and average loudness were found to help explain perceptual judgments of material. These researchers have generally studied material perception for purposes other than for ultimately helping the hearing impaired. This fact may or may not prove to be negligible, but it should be kept in mind when attempting to apply the results to assistive listening devices.

Using synthesized sounds may prove to be useful for studying everyday perception in normal-hearing listeners. However, before making conclusions about the perceptual capabilities of listeners, the synthesized sounds need to be validated with reference to real-world, live sounds. Using inaccurately synthesized sounds will lead to a misconception of how subjects perceive sounds. Even though a researcher may show that perception of a sound may be completely determined by a cue or set of cues, such a result is of little use if the sound is not representative of sounds in the real world.

Information about how normal-hearing listeners perceive size, material, and height could be used, in combination with knowledge concerning how the auditory systems of hearing-impaired subjects are different than normal-hearing subjects (step four), to help explain why the hearing impaired have trouble with perceiving impact sound event properties (step five). However, it should be kept in mind that, because it is not very clear how normal-hearing people perceive these properties, attempts to correct hearing impairment at this stage may be misguided.

The results of the experiment presented in Chapter 5 may provide some insight into the task outlined in step five. The results suggest that adequate audibility may be an important part in the accurate perception of the everyday-sound properties tested. While this hypothesis should be confirmed, it hints that hearing loss as it is classically defined, by an increase in the minimum audible threshold, may help to partially explain poor perception.

The experiments reported in both Chapters 4 and 5 have helped to address step six. It has been mentioned that the results presented in Chapter 4 suggested that spatial cues may be important for the perception of size in normal-hearing listeners. This indicates that it might be beneficial to use care when selecting a type of hearing aid. For example, hearing aids that are better at maintaining spatial cues, such as in-the-canal and completely-in-the-canal hearing aids, may be preferred over behind-the-ear hearing aids, which do not as well maintain the influence of the listener's pinna. Additionally, the results of the experiments presented in Chapter 4 may suggest that it would be best to make sure that both ears of a hearing-impaired listener are treated when their hearing disorder is treated. Unilateral hearing aid fittings for treatable bilateral losses may result in poor spatial hearing and therefore a worsened ability to perceive size.

By testing normal-hearing listeners and hearing-impaired listeners with and without their hearing aids, the experiment reported in Chapter 5 has helped to also address step six. It provided evidence that the hearing aids tested were not making perception of material, length, or height worse. In some cases, the hearing aids even helped to improve the perception of these three characteristics. This does not mean that the hearing aids have not hurt perception in some way, but if they have, they have also aided it in ways that are at least compensatory. A great deal of further research is likely required to determine the best ways for hearing aids to improve the perception of material, length, height, and other ecologically relevant parameters. The results presented here are only the very start of this task.

6.3 Experimental Methods

It is not certain that the results measured in the listening test environment of these experiments are necessarily valid in the real world. The results could be underestimates or overestimates of real-world performance, and cues found to be relevant in these experiments may not be cues used in perfectly natural listening situations (Handel, 1989). Although many attempts were made to make the test situation as natural as possible for the test subjects, the experimental nature of the situation was certainly an obvious part of the listening test for all subjects. Subjects put in unnatural situations may respond differently than they would in real life. For example, they may become excessively analytical.

In their everyday lives, people generally make perceptual judgments without hesitation and without conscious thought. Only when forced to recall a past event do people usually spend time analyzing a situation. It has been shown that by preventing subjects from becoming too analytical, task performance increases (Heft, 1993). An attempt to do this was made in the experiment described in Chapter 4 by requiring subjects to respond within a short amount of time. Limiting the reasoning of the subject could also be accomplished by making the task of interest to the experimenter a secondary task from the subjects' perspective. It appears that the ecological validity of a perceptual judgment may be increased by reducing the possibilities for a subject to become analytical while performing the task or making the perceptual judgment. Further investigations concerning this idea would be helpful for the design of future experiments.

One way of creating an experiment in which the subjects are less focussed on the auditory stimuli may be to involve multiple perceptual systems. For example, by presenting stimuli that produce both visual and auditory signals and then asking a subject to make visual judgments about a particular attribute of the stimuli, controlled variations in the auditory information provided to subjects may be performed and the influence of the judgments correlated to this auditory information. The experimenter, realizing that the task is both auditory and visual in nature, may take advantage of the fact that the subject believes the task to be a primarily visual one (e.g., size perception). This sort of McGurk (McGurk & MacDonald, 1976) effect for everyday sounds has also been suggested by others (e.g., Saldaña & Rosenblum, 1993; Rosenblum, 2004).

Conducting ecologically relevant research may require that the perceptual systems are not treated separately, but that they are considered together. Just as for speech, in which auditory and visual information are both relevant to perception, treating the sensory systems separately may not produce results that are optimally applicable to the real world. This idea is probably also relevant to the perception of everyday sounds. Researchers who wish to help solve perceptual problems solely by addressing the ears of a person may still do so, but they must also realize that the ears are influenced by the eyes and even other parts of the rest of the body.

Of direct relevance to the present experiments, it would be interesting to explore how well test subjects are able to make visual judgments of the same parameters that they judged in the experiments described in this thesis. Although there may be more appropriate tasks, a test like one of those presented in Chapters 4 or 5 could be done for auditory, visual, and auditory/visual combined conditions, each using the same response apparatus. The results could help investigators understand the relative abilities of the perceptual systems and the ways in which the systems work together.

Listeners are sensitive to a great number of informational cues in their everyday environments. Classically, many of these potential cues have been excluded from listening test situations for the sake of simplicity. However, such decisions have been made at the cost of the ecological validity of the experiments. As previously mentioned, the experiments conducted in this thesis have first demonstrated that a classical acoustic stimulus presentation technique, the monophonic recording and playback method, results in perceptual judgments that are inferior to those obtained with more real-world-like listening. The more life-like presentation of live stimuli was then further used in experiments with hearing-impaired subjects. As a group, hearing-impaired subjects with mild to moderate hearing losses performed just as well as normal-hearing subjects. While their audiograms may report that their hearing was not as good as the hearing of the normal-hearing subjects, the hearing-impaired subjects apparently received enough information from the stimuli to make equivalent perceptual judgments. However, if the listeners had been deprived of some of that information, for example through the presentation of monophonic recordings instead of live presentation, it is unclear if their remaining hearing abilities would have been evident.

In general, the questions asked of the subjects in the experiments described in Chapters 4 and 5 were motivated by consideration for how people use their hearing in their everyday lives. However, not all of the questions were asked in a way that made the response clearly relevant or of interest to the test subject. It is likely that more accurate information could be attained about the perceptual abilities of the subjects by asking questions that are more directly relevant to them. Modeled after the experimental method used by Carello et al. (1998), an attempt to do this was made in the experiment described in Chapter 4. The subjects positioned a moveable surface in such a way that a rod could just reach the moveable surface if extended from a fixed reference surface. The task was therefore given ecological significance to the subject by indirectly asking if the rod, as a tool, could reach a target. Just as people need to make decisions based on the sounds they hear in the real world, test situations in which subjects are asked to make similar decisions can improve the real-world relevance of the test results.

The results from these experiments have produced basic knowledge of the abilities of normal-hearing and hearing-impaired listeners in ecologically relevant listening tasks. However, the results are only the start of a potentially long road to understanding *how* listeners perceive the event attributes that are so relevant to their daily lives. Ecologically motivated experimental techniques are only beginning to see the light of day in auditory research laboratories, and it is likely that these fundamental tools need improvement at the same time that knowledge of real-world hearing grows. Determining which questions to ask may be a large part in finding useful answers.

6.4 Future Work

The findings described in Chapter 4 should be validated, and a comparison of the influence of stimulus presentation method for other types of everyday sounds should be conducted. Performance differences resulting from various presentation methods may be larger or smaller than those observed in this project, depending on the types of stimuli being presented and the task of the test subject in the listening test.

In continuation of the work presented in Chapter 4, multi-channel surround sound techniques would also be interesting to investigate. A reproduction technique such as Ambisonics (e.g., Fellgett, 1974; Guastavino et al., 2005), which attempts to create an accurate representation of the recorded sound field over a fairly large area, could free the listener of the need to wear headphones as is more or less a requirement for the binaural technique. Using a reproduction technique utilizing a loudspeaker-based surround sound system could increase the comfort of the test subject and improve the naturalness of the listening situation, while at the same time producing a sound field that does not move when the listener moves his or her head. Such a technique may not be able to reproduce a sound field at the eardrums of a listener as accurately as the binaural technique, but its other advantages may be more valuable in some listening situations. Tests of these kinds of reproduction techniques can help answer these questions.

The results of Chapter 4 suggest that spatial cues may be an important factor in size perception, but confirmation of this result and perhaps a test designed to answer this question directly is needed in order to make this claim with more certainty. Such a test could involve making binaural recordings with an acoustic manikin, evaluating test subject performance when the equalized recordings are presented to test subjects dichotically, and then evaluating test subject performance when presenting a single "ear" of the recordings diotically. An experiment conducted in this way would provide further evidence that recording fidelity issues, other than sacrificed spatial information, could not be responsible for the decrease in performance observed in the experiments of Chapter 4. Presenting the "left" ear signal, for example, to both the left and right ears of the test subject, will result in elimination of the interaural time and level differences that are important to localization in the horizontal plane (Blauert, 1996). Subjects would still have some spatial information with this setup (Batteau, 1967), but a large part of it would be eliminated. If these results demonstrate a reduction in test subject performance, further support will be supplied for the hypothesis that spatial cues contribute to auditory length perception.

It would also be interesting to ask listeners about length perception using live stimulus presentation in an anechoic chamber compared to a normal room. A comparison of results in anechoic and reverberant environments could provide information concerning whether or not the influence of reflections aids or hinders the perception of length. In theory, reflections may provide further information about the shapes, sizes, and orientations of objects involved in sound events, but whether or not this information can be utilized by listeners is unclear.

Although the influence of presentation method was demonstrated in Chapter 4 for normal-hearing subjects, it would also be valuable to investigate performance differences between presentation methods for hearing-impaired subjects. Even with a signal of raised level, hearing-impaired people may not be as good as normal-hearing people at extracting invariants from a sound that is lacking in fidelity. Although there was no performance difference found for normal-hearing listeners between live stimuli presentation and the binaural playback of stimuli recorded with an acoustic manikin, it is possible that a difference exists for those suffering from hearing impairment. Likewise, it is feasible that the performance difference found for normal-hearing subjects could be even greater for hearing-impaired subjects when listening to live stimuli presentation compared to stimuli presented from monophonic recordings.

If either of these hypotheses are true, it would be even more important to use an accurate stimulus presentation technique when testing with hearing-impaired subjects, as opposed to the normal-hearing subjects that were used in the experiment presented in Chapter 4. On the other hand, experiments may alternatively show that the hearing impaired are unable to differentiate between monophonic recordings and live presentation. These ideas should be investigated prior to assuming that the findings of Chapter 4 apply to all listeners. In either case, live presentation is certainly the safest method if it is a concern to represent sound events in a perceptually valid manner. If an understanding of real-world perception is of interest, other techniques should be validated prior to their being used in listening tests of which the results will be used to explain perception.

In Chapter 5, the results for the entire group of hearing-impaired subjects indicated that their hearing aids helped them to perceive length and height more accurately. Only material perception was not improved. Normal-hearing test subjects were no better or worse when equipped with their hearing aids than without hearing aids. The hearing aids have improved the perception of some properties for the hearing-impaired subjects, and have at least not caused a decrease in the abilities of either of the subject groups to perceive the properties of the impact sound events.

Future investigations of whether or not audibility is responsible for the improvement observed by the complete aided hearing-impaired group could be conducted. To test this, the amount of amplification in a hearing aid could be varied, and perceptual judgments of subjects checked at various presentation levels. Test subjects and stimuli would ideally be selected so that the subjects were unable to hear the stimuli without hearing aids. If performance could be increased to a level equivalent to that of normal-hearing subjects, while all other hearing aid parameters other than gain were held constant, it could be stated with more certainty that audibility was responsible for the results.

The fact that the group of hearing-impaired subjects suffering from only mild to moderate losses performed equally as well as their normal-hearing counterparts was somewhat unexpected. When interpreting this result, it is important to keep in mind a few issues if one wishes to generalize the findings. The result may reflect that the test conditions were not difficult enough to elicit differences between the groups, or that these hearing-impaired subjects truly have no more trouble than normal-hearing listeners at perceiving the three parameters investigated. Tests could be performed to investigate this hypothesis.

In order to mask soft background sounds and to provide a stable input level for the hearing aid compression systems to return to between stimulus presentations, a fan was used as a noise source in the test environment. It was expected that this fan might also serve to make the task more difficult for the mild to moderately hearing-impaired subjects than for the normal-hearing subjects, but the disturbance was apparently not great enough to reduce task performance for one group but not the other. It is possible that a more complicated acoustic environment, with louder distractors, more distractors, distractors placed at different positions, or distractors of a different type, could evoke differences between normal-hearing and hearing-impaired subjects. Similarly, the use of quieter target stimuli may present a more difficult task for hearing-impaired listeners than for normal-hearing listeners. Future experiments could examine such issues.

It should also be kept in mind that there exists a wide variety of hearing losses. Beyond the shape of a hearing-impaired subject's audiogram, there are likely other important characteristics of the hearing-impaired subject's ability to hear that contribute to the perception of material, length, and height. Beyond the test of impact sound perception, the sensitivities of the hearing systems of the subjects who participated in this experiment were only tested by means of an air-conduction audiogram. This measurement is the standard way of categorizing a person as "hearing impaired", but it may be a poor or incomplete measure of the abilities necessary to hear everyday sounds. Other functional measures of the sensitivity of the ear, such as the ability to discriminate between frequencies, may provide useful alternatives for correlating real-world perception to hearing loss.

A test that measures information transmission instead of sensation thresholds may ultimately prove to be the best test of hearing ability. A classification system for categorizing hearing losses could be based on an information-based test (e.g., tests of speech intelligibility and everyday sound perception), as opposed to a measure of the minimum audible threshold. Such a test would of course need to be suitably general, but the idea of basing the measurement off real-world listening tasks is desirable when compared to a test based on the sensations elicited by sounds (i.e., pure tone audiometry). Audiometric measurements, which are based on testing tones amplified to various levels, may be attractive simply because assistive listening devices are classically designed to provide support using similar amplification techniques.

As mentioned in Chapter 2, it has previously been shown that visual feedback improved the consistency and accuracy of rod length estimation for normal-hearing subjects (Wagman, 2003a). It would be interesting to determine if feedback also helped hearing-impaired subjects. Such knowledge could suggest general techniques for rehabilitating those suffering from hearing impairments.

In general, all of the findings from this project should be verified with additional test subjects. Of particular importance may be the experiments with hearing-impaired listeners. Making accurate conclusions for large groups of people may require many more subjects and subjects who are representative of many different levels of hearing ability. This point was illustrated in Chapter 5 where it was shown that the inclusion of just two severely impaired subjects resulted in a different outcome for many of the statistical tests that were performed. On the other hand, this fact should also serve to point out that there are a wide variety of hearing losses and problems associated with hearing losses. Grouping subjects into only two groups of “normal hearing” and “hearing impaired” may be too coarse of a categorization system for performing investigations that can effectively lead to treatments for hearing difficulties. Attempting to make conclusions for large categories of people may be inappropriate, particularly so when considering the fact that the audiogram also appears in many cases to be a poor descriptor of an ability to hear in the real world. Future experiments testing the perception of a large number of hearing-impaired subjects’ abilities to hear everyday sounds may even help to reveal more appropriate ways of grouping subjects – perhaps by some other metric than the air-conduction audiogram.

The experiments described in Chapters 4 and 5 have focused on understanding *what* people are capable of perceiving from sound, but have not focused on attempting to explain perception. It would obviously be of great benefit to ultimately be able to predict and explain the perceptual judgments observed in these experiments and others. Carello et al. (1998) attempted to correlate the results from a simple acoustic analysis of rods being dropped on a floor to perceptual results for subjects’ estimates of the lengths of the rods, but a more sophisticated analysis may prove revealing. The analysis of Carello et al. involved only simple regressions of perceived length onto signal duration, amplitude, and frequency centroid, and it failed to find parameters that accounted for perception any better than actual length. Multiple regressions involving more acoustic variables may prove fruitful.

6.5 Conclusion

Every object in the world, involved in every impact that ever occurs, has a size and a material property. Initial research produced by the experiments described in this thesis, and in the publications of others, suggests that material and size are often readily perceivable. Because of the prevalence and relevance of material and size as properties specifying objects, a reduction in ability to perceive material or size surely has an undesirable effect. Hearing-impaired listeners may suffer from a general difficulty to identify the objects involved in sound events, a difficulty that they may or may not be able to articulate. A person suffering from hearing impairment may be likely to suggest simply that he or she has trouble hearing what is happening, as opposed to a difficulty in identifying the size and material of objects involved in sound events. This may be similar to the way in which a hearing-impaired person may complain of a difficulty understanding speech, as opposed to a difficulty identifying the vocal-tract articulatory features responsible for consonant production in a speaking person.

While the layperson may not normally consider why he or she has trouble identifying events in his or her surroundings, the experimenter must do so if it is ultimately desired to improve the experience of those suffering from abnormal perception. A consideration of the ecologically relevant information responsible for identifying sound events is clearly a requisite for achieving this goal. It is up to the researcher to conduct intelligent questioning of subjects or to make intelligent guesses at what might help the hearing impaired to hear "what is happening." Furthermore, it is clear that if it is desired to understand the everyday use of the human auditory system, experimenters must find ways to optimize both the ecological validity of experiments and the information provided to the experimenters by the experiments. Although it does not appear to be the approach used in classical psychoacoustics experiments, many indicators suggest that both of these goals may be reached simultaneously.

Appendix A

Sequential Streaming Effects on Pitch Perception

The following project was done in close cooperation with Torsten Dau of the Centre for Applied Hearing Research at the Technical University of Denmark.

Introduction

Spectral, temporal, and spatial auditory cues have been found to be utilized by humans to group and separate sounds in their environments (e.g., Bregman, 1990). Such skills allow listeners to make sense of information contained in complicated acoustic signals by allowing the listeners to focus on only the parts of interest (Cusack & Carlyon, 2004). The disturbance potentially caused by background noise sources and sources to which the listener does not devote attention can be reduced if a listener is able to perceptually separate sources. An understanding of how such auditory events are identified and heard as separate auditory streams is useful for the treatment of hearing disorders (Buchler, Allegro, Launer & Dillier, 2005) and in computational automatic source segregation such as systems designed for speech, environmental sound, and other sound recognition purposes (e.g., Green, 2004; Janku, 2004; Haykin & Chen, 2005).

The relative strengths of auditory grouping cues can be studied by examining situations in which multiple cues are set into competition with one another. In the presence of multiple sound sources, whether or not a particular sound element belongs to one stream or another may vary depending on the precise conditions. A grouping cue that is effective under one set of conditions may not be effective under slightly different conditions or when a listener is performing a different listening task (Darwin, Hukin & Al-Khatib, 1995). If a particular grouping cue is stronger than an alternative grouping cue, that sound element may be drawn into the stream created by the dominant cue.

Regularities in the spectrum of a signal, such as those that may be present in a note produced by a musical instrument, can lead to the perception of a single acoustic source. At a simplified level, such a note can be thought of as being formed from multiple, harmonically related tones (Rossing, 1990). If one of these harmonics is slightly mistuned from its correct frequency, the overall perceived pitch may be altered. As the amount of mistuning increases, the influence of the mistuning on the "residue pitch" of the complex tone eventually decreases. Mistunings above around ±3% result in a reduced influence of the mistuned component (Moore, Glasberg & Peters, 1985). With sufficient mistuning beyond this point, the harmonic may be heard separately. At this point, its presence is then not as important in the perception of the residue pitch of the complex tone (Darwin et al., 1995). In terms of a harmonic sieve pitch perception model such as that by Duifhuis, Willems & Sluyter (1982), it has been found that a harmonic, which has been mistuned by up to ±3%, should still be fully included in the calculation of residue pitch. Mistuning the harmonic by larger amounts leads to a reduction in the influence of the harmonic, and shifts of ±6 to 8% are enough to eliminate measurable effects of the harmonic (Moore et al., 1985).

Research described by Moore et al. (1985) found that the fourth harmonic of a twelve-harmonic complex tone was on average the most influential harmonic in the pitch of the complex tone. Such a mistuned harmonic may affect the pitch of the complex tone when the tone is heard in isolation, but evidence exists indicating that the influence of preceding tones can reduce the effect of the mistuned harmonic. Specifically, evidence for the influence of sequential streaming on pitch perception has been provided by Darwin et al. (1995) who showed that a series of four preceding pure tones ("precursors") at the frequency of a mistuned fourth harmonic were effective at reducing the influence of the mistuned harmonic on the residue pitch of the complex tone. Pitch shifts measured with the preceding tones were smaller than pitch shifts measured with the complex tone in isolation. This suggested that sequential streaming was responsible for the mistuned harmonic being captured by the precursors. However, peripherally based mechanisms of neural adaptation could also have been at least partly responsible for reducing the influence of the mistuned harmonic.

If a *following* tonal sequence was found to be as effective as a preceding tonal sequence at removing the perceptual influence of the mistuned harmonic, then further evidence against adaptation and for sequential streaming being responsible for the results presented by Darwin et al. (1995) would be provided. Researchers studying everyday sound producing events have shown that how an event unfolds can influence the perception of how it began (Fowler, 1990). In more traditional psychoacoustic experiments, it has been previously noted that there can be retroactive perceptual consequences for activity occurring after a stimulus (e.g., Darwin, 1984a; Roberts & Moore, 1991), that grouping can operate retroactively (Darwin, 1984b), and of even greater relevance that following tones can affect the perceived sound of a vowel (Darwin, Pattison & Gardner, 1989). Following tonal sequences ("postcursors") have also been shown to have similar effects in certain signal detection tasks (Dau, Ewert & Oxenham, 2004), using the same parameters for the sequences as here. In this case, backward masking could be ruled out due to the timescales and levels involved (Oxenham & Moore, 1994; Elliott, 1971), and the possibility of peripheral adaptation could be discarded because it is known that a peripheral neural response is not affected by trailing stimuli (Kiang, Watanabe, Thomas & Clark, 1965). An attempt was made here to test the idea of whether or not postcursors were effective at removing the influence of a mistuned harmonic on the residue pitch of a complex tone. Therefore, the experiment was one in which sequential grouping cues were put into competition with simultaneous grouping cues. A pitch comparison technique was used in order to test the hypothesis. Stimuli were presented in a similar manner as in Experiment 1 of Darwin et al. (1995), but with the option of testing postcursors in addition to precursors.

Methods

Six volunteer subjects, including the author, participated in the experiment. All were recently found to have normal hearing and were highly experienced with psychoacoustic tests but mostly inexperienced with pitch matching experiments. Five of the subjects had some musical training. After completing an approximately half-hour-long training session to make sure the subjects were able to compare the pitches of complex tones, subjects completed two distinct portions of a listening test.

In both conditions, subjects heard two 90 ms twelve-component complex tones separated by 500 ms of silence. In the "isolated" condition, the target tone had a mistuned fourth harmonic and a second, adjustable complex tone had a fundamental frequency that was varied for comparison to the target. In the "postcursor" condition, the target tone was immediately followed by four 90 ms pure tones at the same frequency as the target's mistuned harmonic. Each of these pure tones was separated from one another, and from the target, by 50 ms. The adjustable

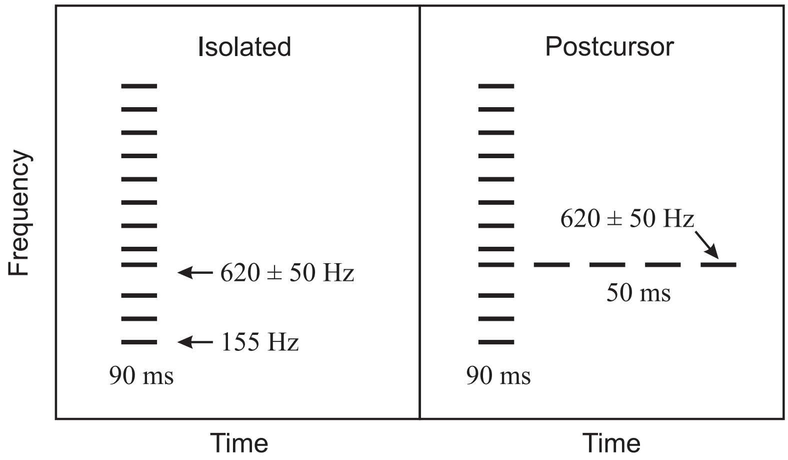


Figure A.1: Target stimuli for isolated (left) and postcursor (right) conditions, each with a variable fourth harmonic.

tone was presented prior to the target in this condition. A diagram of the target stimuli is shown in Figure A.1. The fundamental frequency of the target was always 155 Hz, and only the fourth harmonic was mistuned. The second through twelfth harmonics of the tone with variable fundamental were allowed to change according to a normal harmonic ratio. Stimuli were generated at a sampling rate of 48 kHz, and all tones had Hanning-window shaped onsets and offsets of 5 ms duration. There was a one-second pause between trials.

For comparison to the earlier work of Darwin et al. (1995), four of the test subjects also completed a "precursor" condition. In the precursor stimulus, the captor tones preceded the complex tone, just like that in Experiment 1 of Darwin et al. (1995). The precursor stimulus was therefore a time-reversed version of the postcursor stimulus shown in Figure A.1. In this precursor condition, the adjustable tone followed the target tone after the usual 500 ms of silence.

The test was controlled by an adaptive, non-interleaved simple up-down method in which the response of the test subject, when asked to compare the pitch of two tones, determined whether the fundamental of the adjustable complex tone was increased or decreased. The fundamental began at a random frequency and was allowed to vary between the limits of 155 ± 6 Hz. The step size was progressively decreased from 2 Hz to a smallest step size of 0.25 Hz. The mean frequency of the last six reversals at the minimum step size was then taken as the measured residue pitch.

The fourth harmonic of the target tone was mistuned from its natural harmonic value of 620 Hz by -8%, -3.2%, -1.6%, 0%, +1.6%, +3.2%, and +8%, corresponding to 570 Hz, 600 Hz, 610 Hz, 620 Hz (in tune), 630 Hz, 640 Hz, and 670 Hz. Four repetitions of the pitch match were measured for each of the seven mistuning values, making a total of 28 measurement sequences for each of the isolated and postcursor conditions. Half of the test subjects first completed the isolated condition, while the other half started with the postcursor condition. Subjects who also completed the precursor condition did so after completing both the isolated and postcursor conditions. Subjects were allowed to complete the test at their own pace, sometimes over a few days, but were asked to take periodic breaks at minimum. A total of 56 measurement sequences were ultimately collected for each subject, requiring approximately two hours of test time. The collection of precursor data added 28 measurement sequences and about one hour to the test.

Stimuli were presented to the left ear of a pair of Sennheiser HD 414 headphones. These headphones were chosen because of their use in the study presented by Darwin et al. (1995). A Brüel & Kjær PULSE analyzer, a Type 4153 IEC 318 compliant artificial ear, and a Type 4230 sound level calibrator were used to check and calibrate the stimuli. The level was set so that a single 1000 Hz tone, of identical electrical amplitude to all other harmonics used in the test, produced a sound pressure level of 60 dB. The headphones were not equalized, but the frequency response was flat to within ±3 dB across the frequency range of interest.

Results

As proposed by Darwin et al. (1995), data from each subject were normalized by expressing the subjects' average pitch matches as deviations from their average pitch match when the target had its fourth harmonic in tune (620 Hz). The mean data and standard errors for all subjects are shown for the isolated and postcursor conditions in Figure A.2. Ipsilaterally measured results from the isolated condition of Darwin et al. (1995) are plotted for comparison. For the isolated condition, the general pattern is qualitatively similar to that observed by Darwin et al., but the pitch shifts of interest are much smaller here. Analysis of variance showed that the effect of the amount of mistuning is significant for the isolated/postcursor factor ($F_{6,70} = 8.57$, $p < 0.0001$). However, neither the main effect of the isolated/postcursor factor was significant ($F_{1,70} = 0.50$, $p = 0.48$), nor was the interaction of this factor with amount of mistuning ($F_{6,70} = 0.23$, $p = 0.96$). A difference in the pitch estimates of the isolated and postcursor conditions was therefore not found.

Figure A.3 shows the mean change in matched fundamental averaged for the positive and negative mistunings for all test subjects. Again, the mean data (ipsilateral

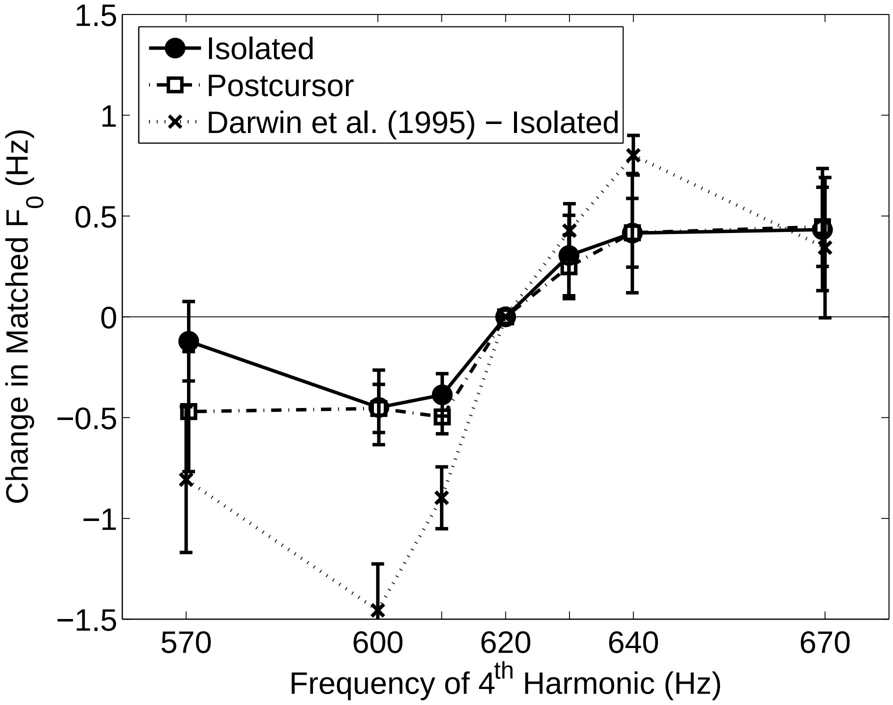


Figure A.2: Mean residue pitch (with standard errors) of a complex tone with mistuned fourth harmonic. Pitch matches for tone in isolation and with postcursors are expressed as the frequency displacement of the matched tone's fundamental. Ipsilaterally measured isolated condition data from Darwin et al. (1995) are shown for comparison.

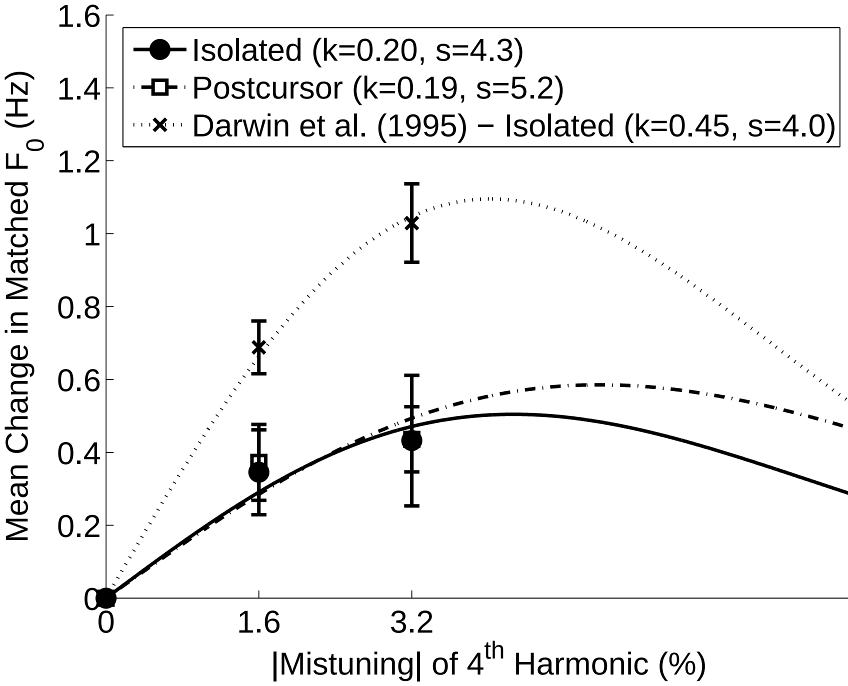


Figure A.3: Average of positive- and negative-mistuning residue pitch (with standard errors) of a complex tone with mistuned fourth harmonic. Pitch matches for tone in isolation and with postcursors are expressed as the frequency displacement of the matched tone's fundamental. Isolated condition data (ipsilateral and contralateral mean) from Darwin et al. (1995) are shown for comparison.

and contralateral average) as measured for the isolated condition by Darwin et al. (1995) have been reproduced for comparison. The means were determined by calculating half the difference between the pitch matches of the corresponding positive and negative mistunings of the fourth harmonic. Neither the main effect of the isolated/postcursor factor was significant ($F_{1,30} = 0.40$, $p = 0.53$), nor was the main effect of the amount of mistuning ($F_{2,30} = 0.18$, $p = 0.84$). Furthermore, the interaction of these two factors was not significant ($F_{2,30} = 0.25$, $p = 0.78$).

As done by Darwin et al., the data in Figure A.3 has been fit with a Gaussian derivative function:

$$\Delta F_0 = k\Delta f \, \exp(-\Delta f^2/2s^2)\,,$$

where k is a scaling factor proportional to the contribution that the mistuned harmonic makes to the pitch of the complex tone, Δf is the amount of mistuning, and s is the standard deviation of the Gaussian function. The values of k for the isolated and postcursor conditions were 0.20 and 0.19, respectively, and the

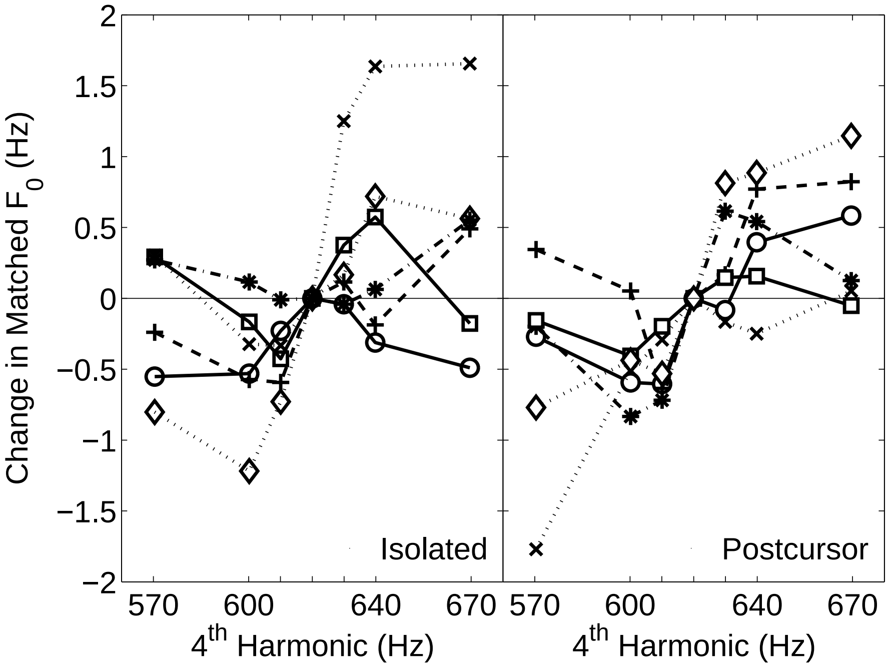


Figure A.4: Individual test subject residue pitches (normalized), for the isolated (left) and postcursor (right) conditions, expressed in terms of frequency displacement from 155 Hz for each possible mistuning of the fourth harmonic. Symbols differentiate data from the six test subjects.

values for s were 4.3% and 5.2%. The same tendency of the mistuned harmonic to elicit a change in residue pitch has been observed, but the pitch shifts were not as large as those reported by Darwin et al. (1995). The value of the width parameter (s) was similar, but Darwin et al. reported a much larger value of the amplitude parameter ($k = 0.45$) for the isolated condition.

For the subpopulation of subjects who also completed the precursor condition, an analysis of variance on the one-sided pitch shift data showed no significant difference between the isolated and precursor conditions ($F_{1,18} = 2.47$, $p = 0.13$). This is in opposition to the trend shown by Darwin et al., in which the precursors resulted in significantly reduced pitch shifts.

Mean pitch matches for the individual test subjects are shown in Figure A.4. It can be seen that variability between test subjects is high, even in the normalized data where the measured pitch shift for the in-tune fourth harmonic (620 Hz) has been subtracted as a standard bias. None of the individual subjects showed a strong effect in the isolated condition with clearly reduced shifts in the sequential streaming condition, as was found in the average data of Darwin et al. (1995).

Discussion

In contrast to Darwin et al. (1995), where an effect of precursors led to significantly reduced pitch shift magnitudes, no effect of postcursors was seen in this experiment. This could be because sequential grouping does not occur or because the grouping differences were simply not measurable with this paradigm. If grouping does not occur, then this does not support the hypothesis that sequential streaming causes the reduction in pitch shifts presented by Darwin et al. (1995). If residue pitch is determined at an earlier stage in the auditory pathway than sequential grouping, then a sequential stream may not include the mistuned component of our target stimulus. Alternatively, postcursors may not be effective if the grouping cue that fuses the complex tone's concurrent frequency components is stronger than the competing sequential grouping cue of the pure tone sequence. In a vowel formant frequency-estimation task, Roberts & Moore (1991) have also found that a captor sequence did not lead to noticeable signs of perceptual grouping. Their results demonstrate that sequential constraints may not be observable in all cases. This seems also to be the case in the conditions investigated here.

It was expected that the isolated condition measurement would produce results like those reported by Darwin et al. (1995), but the results here are not as pronounced. Instead, the isolated condition shifts measured in the present study appear to be on a similar scale to those reported by Moore et al. (1985), though the stimuli were at slightly different frequencies in that case. While the accuracy of the data measured here or in previously published results is impossible to assess, the variability in the data is at least not notably different from that of Darwin et al. (1995).

Differences that exist between data from the isolated condition measured here and that of Darwin et al. (1995) could be due to differences between the test subjects, differences between the test methods, or because the task may be problematic for investigating this phenomenon. All other parameters were nearly identical. Micheyl, Delhommeau, Perrot & Oxenham (2006) have recently shown that pitch discrimination abilities vary widely among people, in particular between musicians and non-musicians for a set of subjects who did not have any previous psychoacoustical training. The authors found that pitch discrimination thresholds for non-musicians were more than six times larger than the discrimination thresholds for classically trained musicians. Between 4 and 8 h of training were necessary to bring the thresholds of the non-musicians to levels like that of the musicians – far more training than was done in the present experiment. While this point is interesting, it probably cannot explain the differences between the present results and that of Darwin et al. (1995). The description of test subjects who participated in the experiments of Darwin et al. appears to also describe the test subjects used in this experiment. It therefore seems unlikely that test subject differences are responsible for differences between the results.

Data measurement methods could be a source of discrepancies. In contrast to the experiments of Darwin et al. (1995) and Moore et al. (1985), in which a method of adjustment was used, the adaptive simple up-down method employed here did not as directly involve the test subject in manipulating the fundamental frequency of the adjustable complex tone. Rolland, Meyer, Arthur & Rinalducci (2002) have noted that a method of adjustments more greatly involves the test subject and is a more usable technique. It was also found by Rolland et al. (2002) that a method of adjustment produced more accurate results than a method of constant stimuli. The adaptive simple up-down method used in our experiment may be considered to produce results similar to the method of constant stimuli, only in a more efficient manner (Levitt, 1971; Dai, 1995). However, whether or not the visual task results from Rolland et al. (2002) are applicable to our auditory task is uncertain.

A final explanation could simply be that this task or analysis techniques may be problematic for measuring sequential constraints on pitch perception. In a similar study, Moore et al. (1985) found large differences between the pitch matches of test subjects. Such variation between test subjects was also observed here, the result of which could be that whether or not differences exist between test conditions, they may not be measurable. The pitch shift differences that we were attempting to measure are very small compared to the variability in the data, possibly making it difficult to detect the effects of sequential streaming on pitch perception.

Though precursors have previously been shown to effectively remove the influence of a mistuned partial from the pitch of a complex tone, our additional test in search of evidence for sequential streaming has not tested positive. In conclusion, we could not support the hypothesis of sequential constraints in pitch perception, but there is also uncertainty whether or not the task and test method are suitable for investigating this specific auditory phenomenon.

Experiment Details: Sequential Streaming Effects on Pitch Perception

Listening Test Description

In this experiment, you will listen to pairs of complex tones and be asked to compare the pitches of the tones. Pitch can be defined as “that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale (Moore, 2003).” In some cases, the complex tones that you will hear will be preceded or followed by four quieter short beeps, but your concentration and judgment should be based on the two complex tones, attempting to ignore all other tones including those of previous trials. In all cases, the tones should be presented to your left ear only. On each trial, you will be asked to answer the following question:

> *Was the pitch of the SECOND complex tone higher or lower than that of the first complex tone?*

If you believe the pitch of the second complex tone was higher than that of the first, then you should press (or click) “Higher (1)”. If you believe the pitch of the second complex tone was lower than that of the first tone, you should press “Lower (2)”.

Before beginning the test, you will be asked to complete a training session in order to make sure that you understand the task and to give you some training in performing the task. The test and training should be conducted in MATLAB in the **Left CAHR booth**. The entire experiment, including training and test sessions, should take about 2.5 hours. It is therefore recommended that you do not try to complete the test all at once. A break of a few hours between test sessions is recommended. You are similarly encouraged to take short breaks at any time during each session.

Before continuing to start the training and test, please follow the items on the attached *Checklist* to make sure that the system is ready for the test. If you have any questions, please ask.

The training is composed of five parts and should last about 30 minutes.

1. Run the first part of the training session as `training1('XYZ')`, where XYZ are your initials, and follow the onscreen instructions. The other training sessions can then be run as `training2('XYZ')`, `training3('XYZ')`, and so on, through `training5('XYZ')`.
2. Please run all 5 training scripts. You may repeat them as often as you like. In addition to these scripts, if you feel it would help you to improve your accuracy by knowing the answer prior to hearing the stimuli, you can *additionally* run training sessions `training1a('XYZ')` through `training5a('XYZ')`

(note “a” after session number). These correspond in difficulty to training sessions 1-5 (without “a”). If you are having trouble getting correct responses, these additional scripts may help you improve your understanding of the task.

Once you have completed the training and assuming you feel that you understand the task, you are ready to perform the actual experiment. There are two parts, which will take about one hour each, including breaks. You should be told which part you should do first.

cd1. `afc_main('PitchMatchUpDown','XYZ','cd1')`, where “XYZ” are your initials.

cd3. `afc_main('PitchMatchUpDown','XYZ','cd3')` - In this condition, you will hear four quieter short beeps after the two complex tones that you are being asked to compare.

Checklist

1. Don't continue with the test unless you are sure you have a normal audiogram. If you are in doubt, ask to be checked.

Outside the Left CAHR Booth

2. Plug the HB7 Headphone Driver (labeled "AMPLIFIER") into the computer's sound output port. The headphone driver has the connector with two wires leading into one headphone connector with duct tape around it.

Inside the Left CAHR Booth

On the HB7 Headphone Driver

3. Make sure the "POWER" switch of the headphone driver is turned on.
4. Make sure the "AC/DC" switch is set to "AC".
5. Make sure the "DIFF" switch is switched DOWN.
6. Make sure the "Gain (DB)" knob is set to "0".
7. Plug the cord of the white and yellow HD 414 headphones into the "PHONO OUTPUTS" jack.

In the RME DIGI Settings, under the "DIGI96/8 PAD (1)" Tab, under "Analog Output"

8. Make sure the "Attenuation" is set to "0 dB".
9. Make sure the "Volume" sliders are all the way to the top of their scales and read "0.0" under each of them.

Final Steps Before Running the Training and Test

10. Make sure Caps Lock is turned off on the keyboard.
11. Turn on the Num Lock if it is not already on.

12. Run MATLAB by double-clicking on the MATLAB icon on the desktop.
13. Enter “bki” as the login name at the login window.
14. Maximize the MATLAB window.
15. Close the “Workspace” window if open.
16. Close the “Command History” window if open.
17. Close the “Current Directory” window if open.
18. Close the “Launch Pad” window if open.
19. Put on the HD 414 headphones, with the red connector on your right ear.

The computer should now be ready for the experiment.

Appendix B

Experiment Details: The Influence of Stimulus Presentation Method on Auditory Perception of Object Length

Listening Test Instructions

The listening test in which you are about to participate is part of a study used to help understand the hearing abilities of humans. In the experiment, you will hear various rods dropped on the floor of the room in which you are sitting. You will be asked to estimate the length of those rods immediately after each is dropped. You will have only a few seconds to respond.

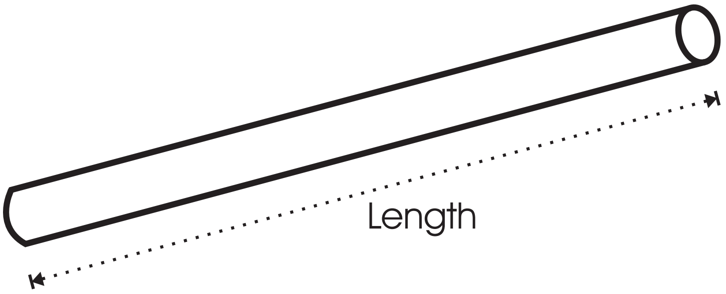
In order to respond, you should position the moveable surface (right) at a point that could just be reached by the rod if it were extended from the fixed reference surface (left). You may move the partition at any time when the red light from the laser is shining on the moveable surface. To notify you that your time is almost up, a series of four beeps will sound (three short beeps followed by one long beep), at the end of which your answer will be recorded. The red light will then go off, and another rod will be dropped. You may be uncertain of your first few responses, but don't worry about this as they will only be used for familiarizing you with the test.

During the test, you are encouraged to move your chair or leave your chair while positioning the moveable surface. This may be physically necessary for an accurate estimation of all rod lengths. It is only asked that you return to the chair and face the wall after you have reported each answer.

The test will be split into four sessions, each lasting approximately ten minutes. In some of these sessions, you will be asked to wear headphones. Your task is identical in this case as in the case when you are not wearing the headphones. You will be allowed a break between each session for as long as you wish.

You will be paid for your participation at a rate of 103.24 kr/hr + 12.90 kr/hr holiday allowance.

If you have any questions now or during the test, simply notify the test operator. You may feel free to stop the test at anytime.



Stimuli

Figures B.1 through B.8 show spectrograms and signal amplitude versus time plots of some of the binaural recordings used as stimuli in the listening test described in Chapter 4. There is one spectrogram and one amplitude versus time plot shown for each of the rod lengths, but five unique recordings of the drops of each rod length were used in the actual test. The signal from the microphone in the right ear of the acoustic manikin is shown. The full recordings used in the listening experiment were 2 s in length, but they have been truncated for presentation purposes here. For reference, Figure B.9 shows the typical background noise level. The plots in Figures B.1 through B.9 have been scaled so that the spectrogram colors can be compared between figures. The maximum decibel levels (relative) in the spectrograms are shown in red and correspond to 0 dBFS. The minimum values, which correspond to -80 dBFS, are shown in dark blue.

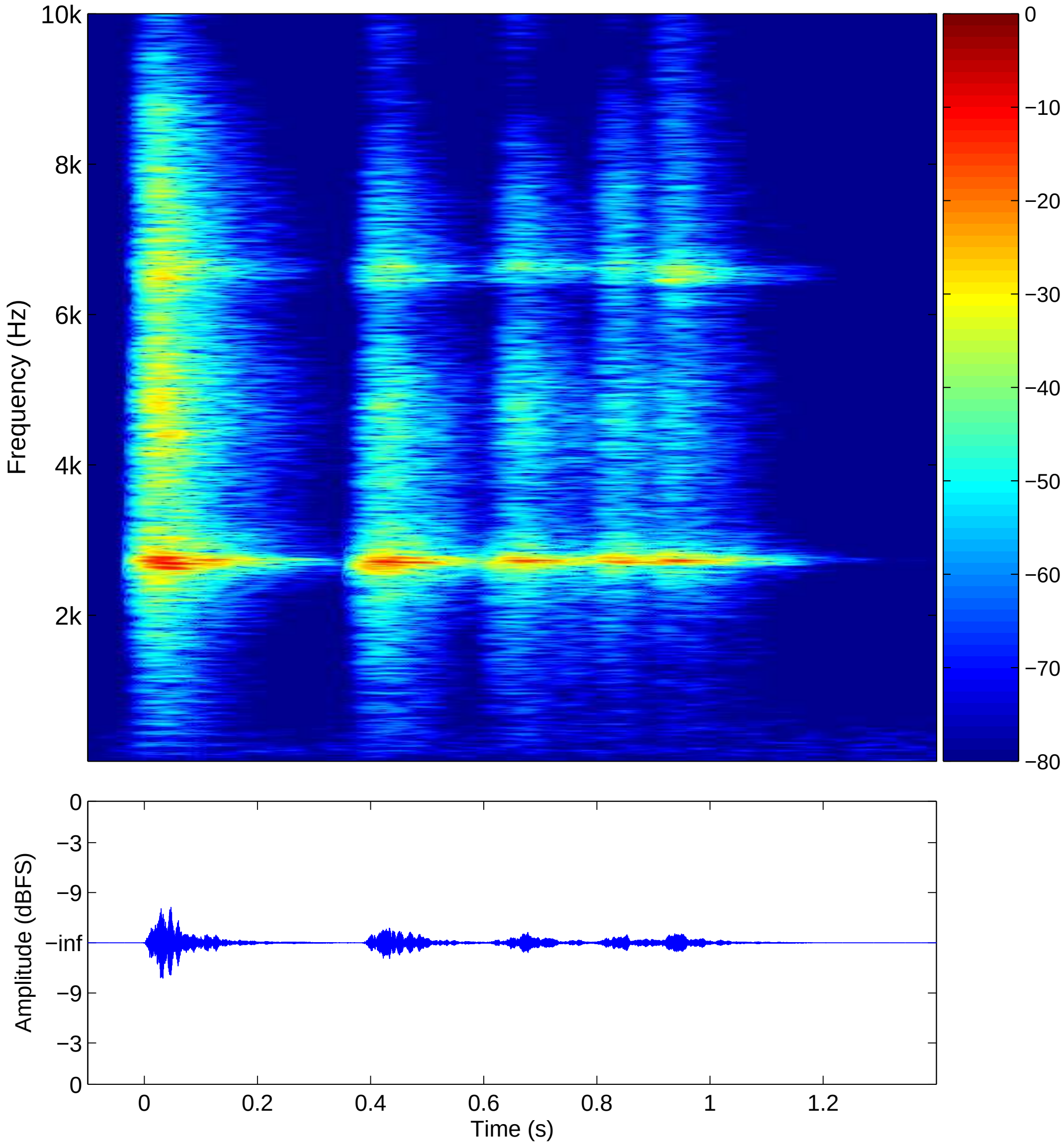


Figure B.1: Wooden dowel, 15 cm long, 13 mm diameter.

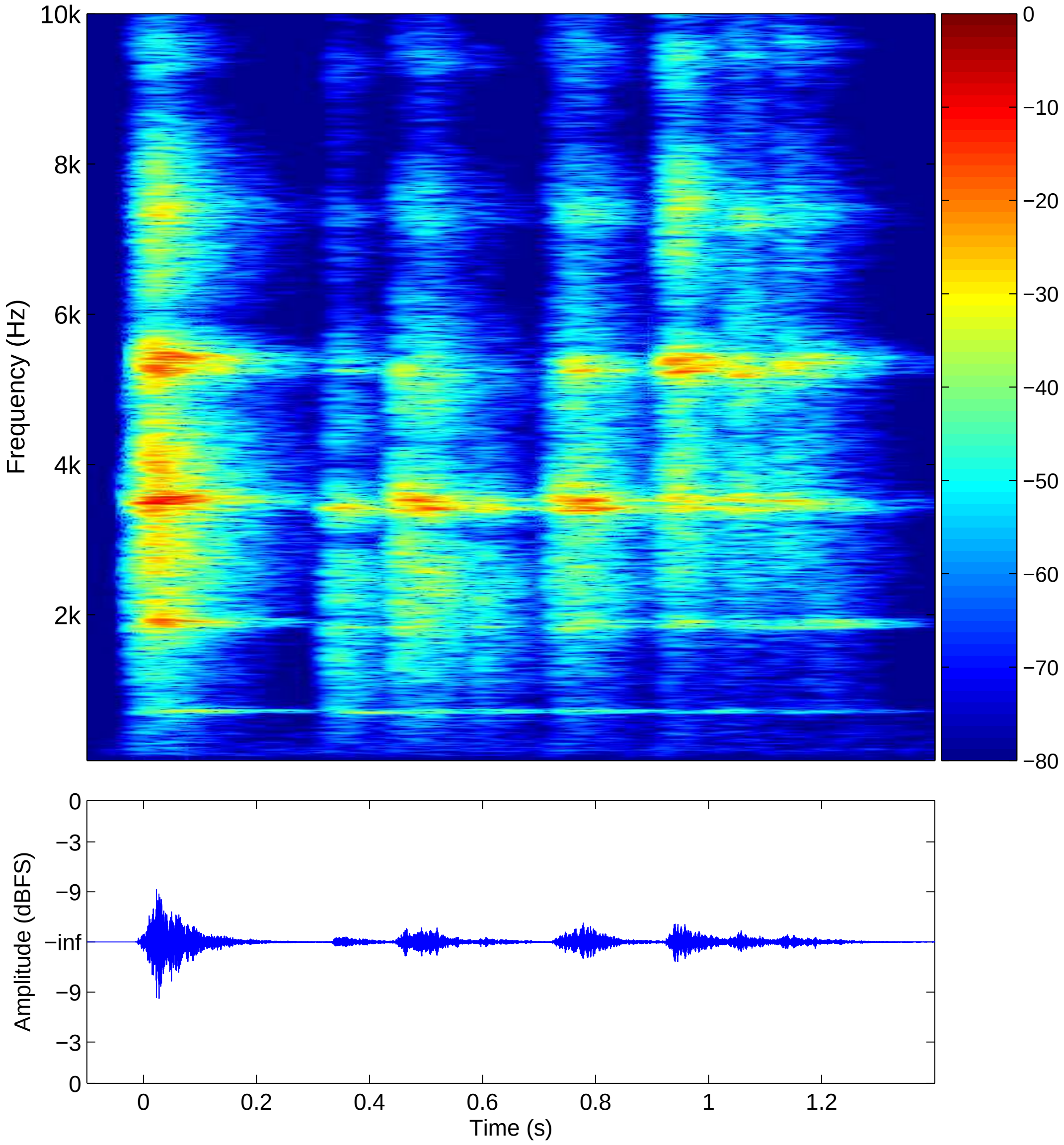


Figure B.2: Wooden dowel, 30 cm long, 13 mm diameter.

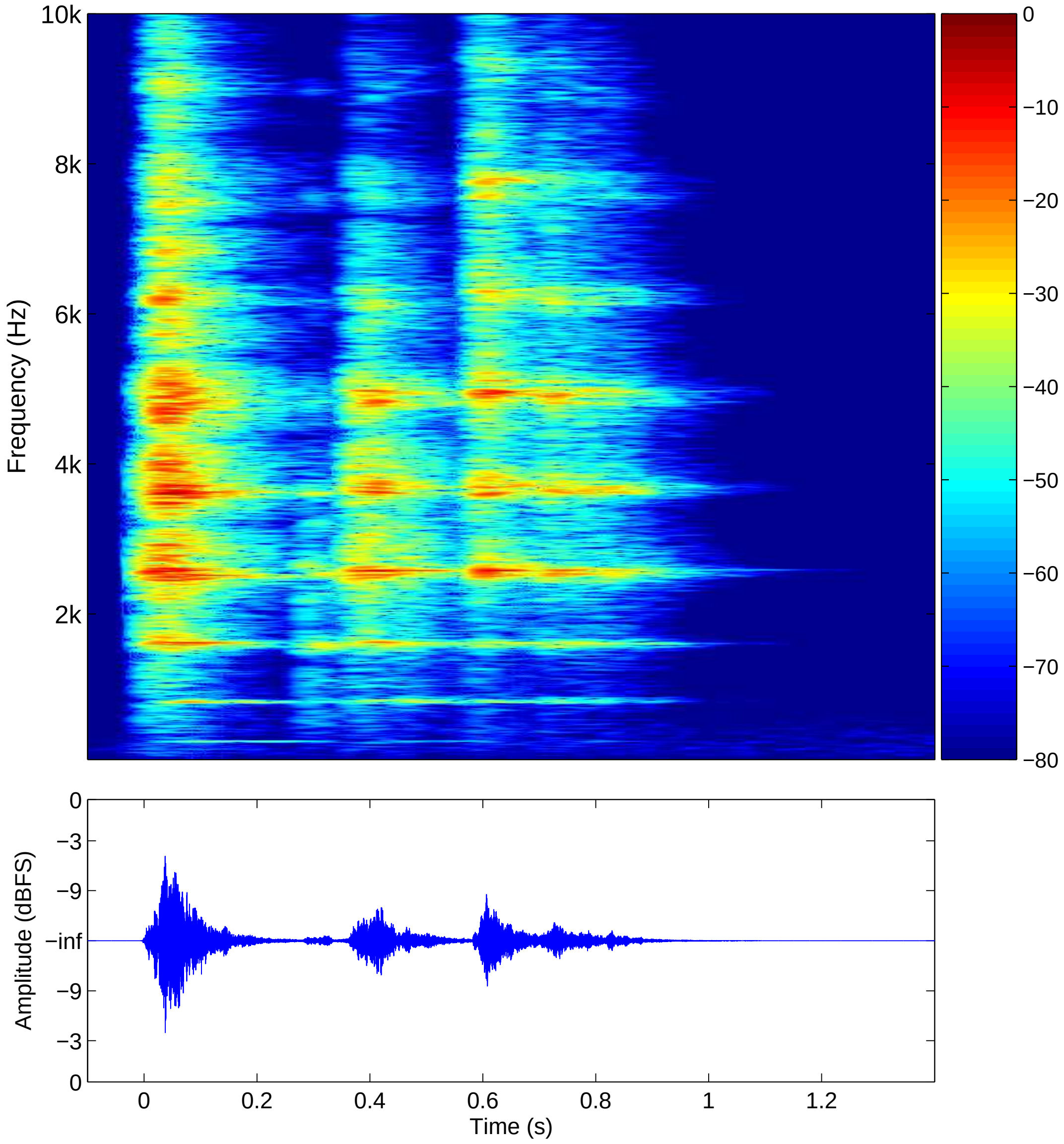


Figure B.3: Wooden dowel, 45 cm long, 13 mm diameter.

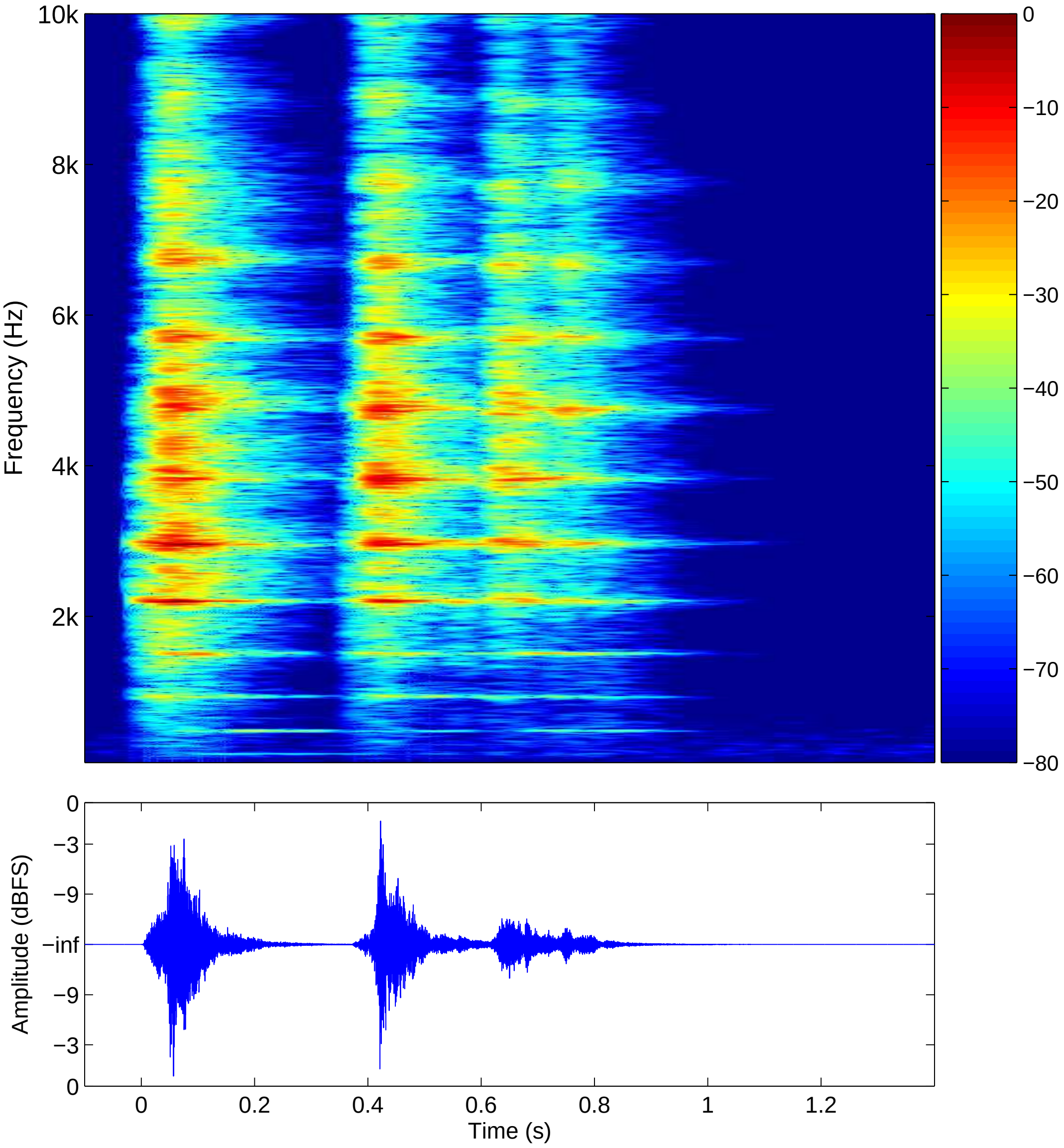


Figure B.4: Wooden dowel, 60 cm long, 13 mm diameter.

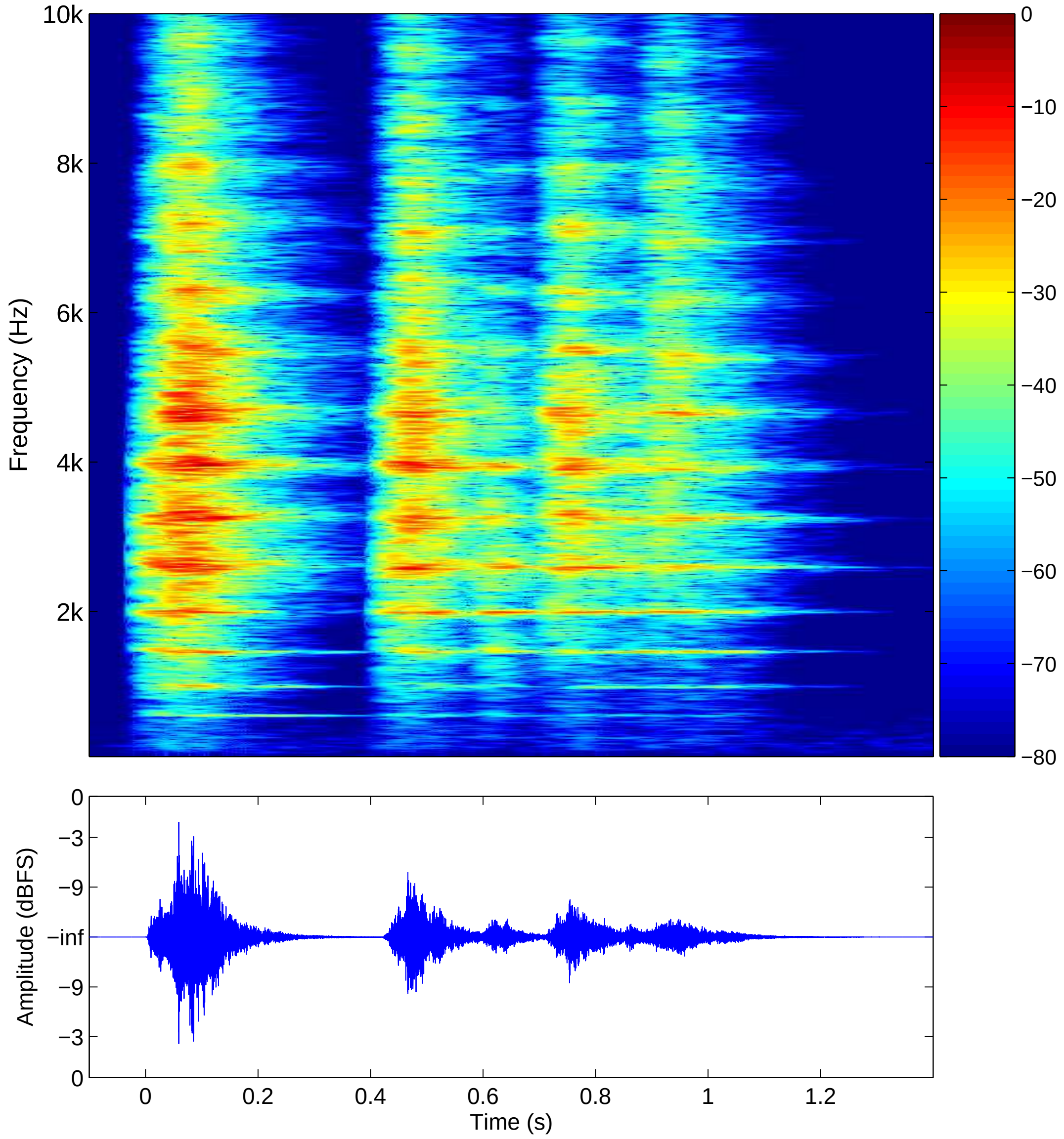


Figure B.5: Wooden dowel, 75 cm long, 13 mm diameter.

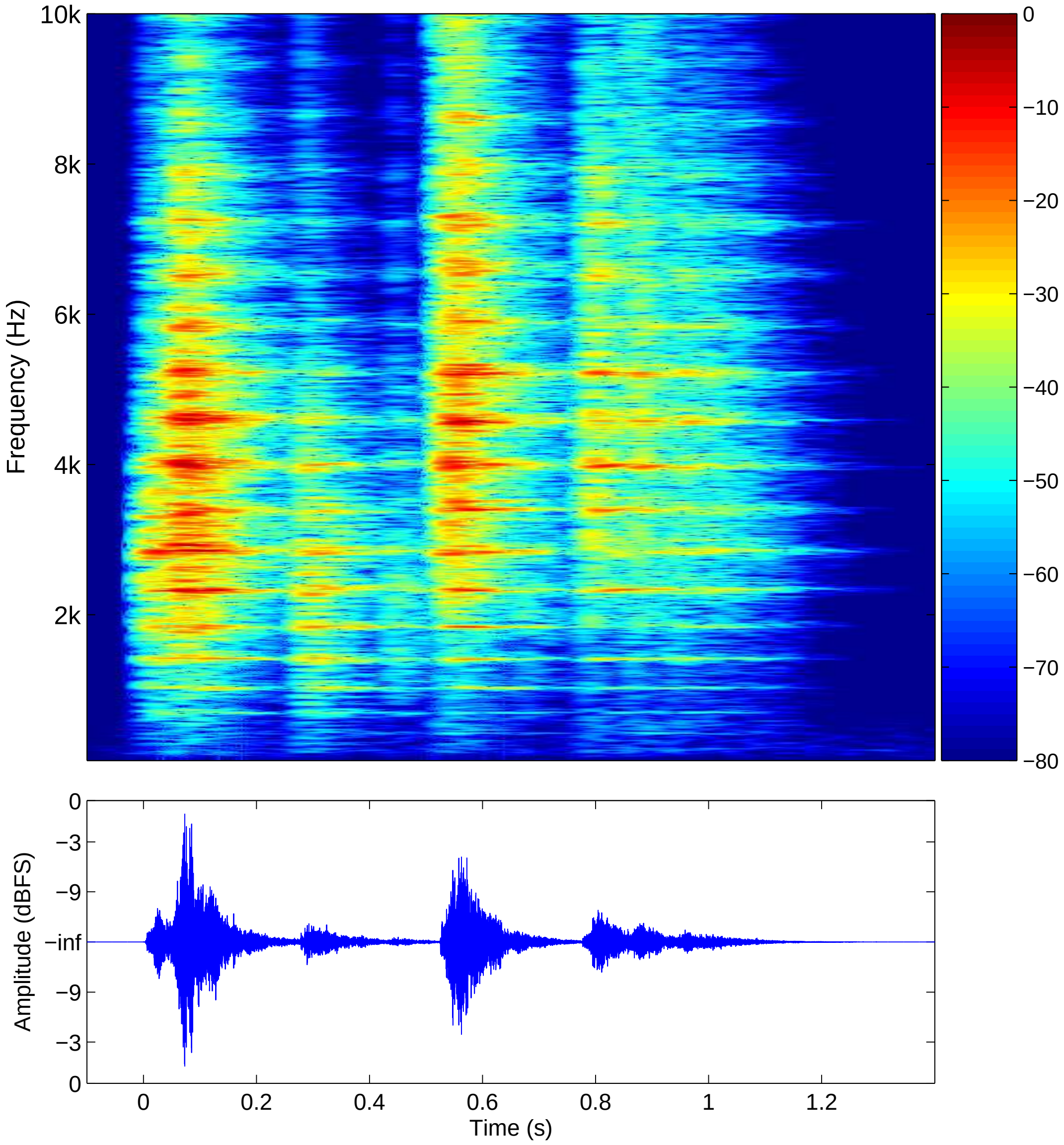


Figure B.6: Wooden dowel, 90 cm long, 13 mm diameter.

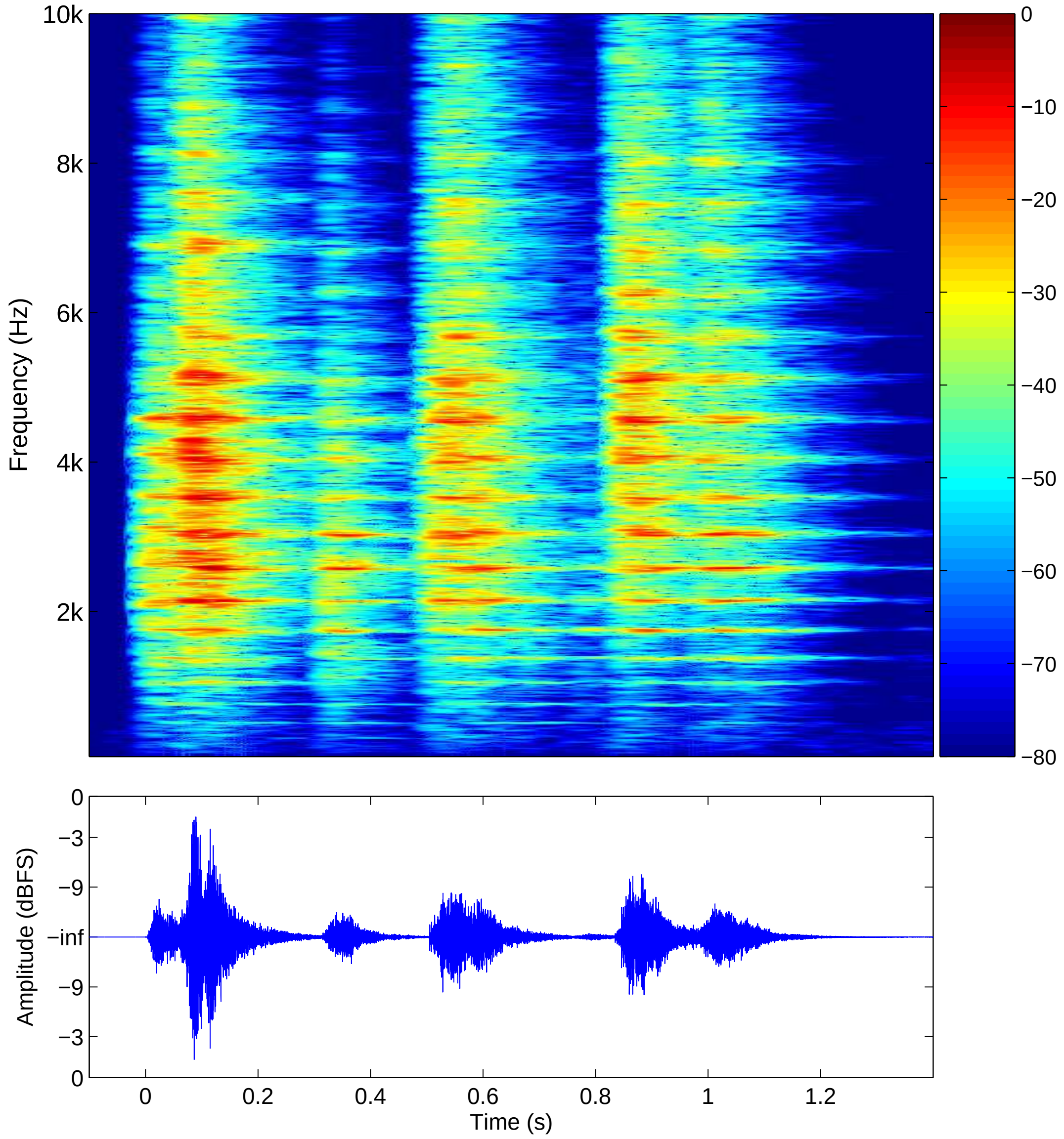


Figure B.7: Wooden dowel, 105 cm long, 13 mm diameter.

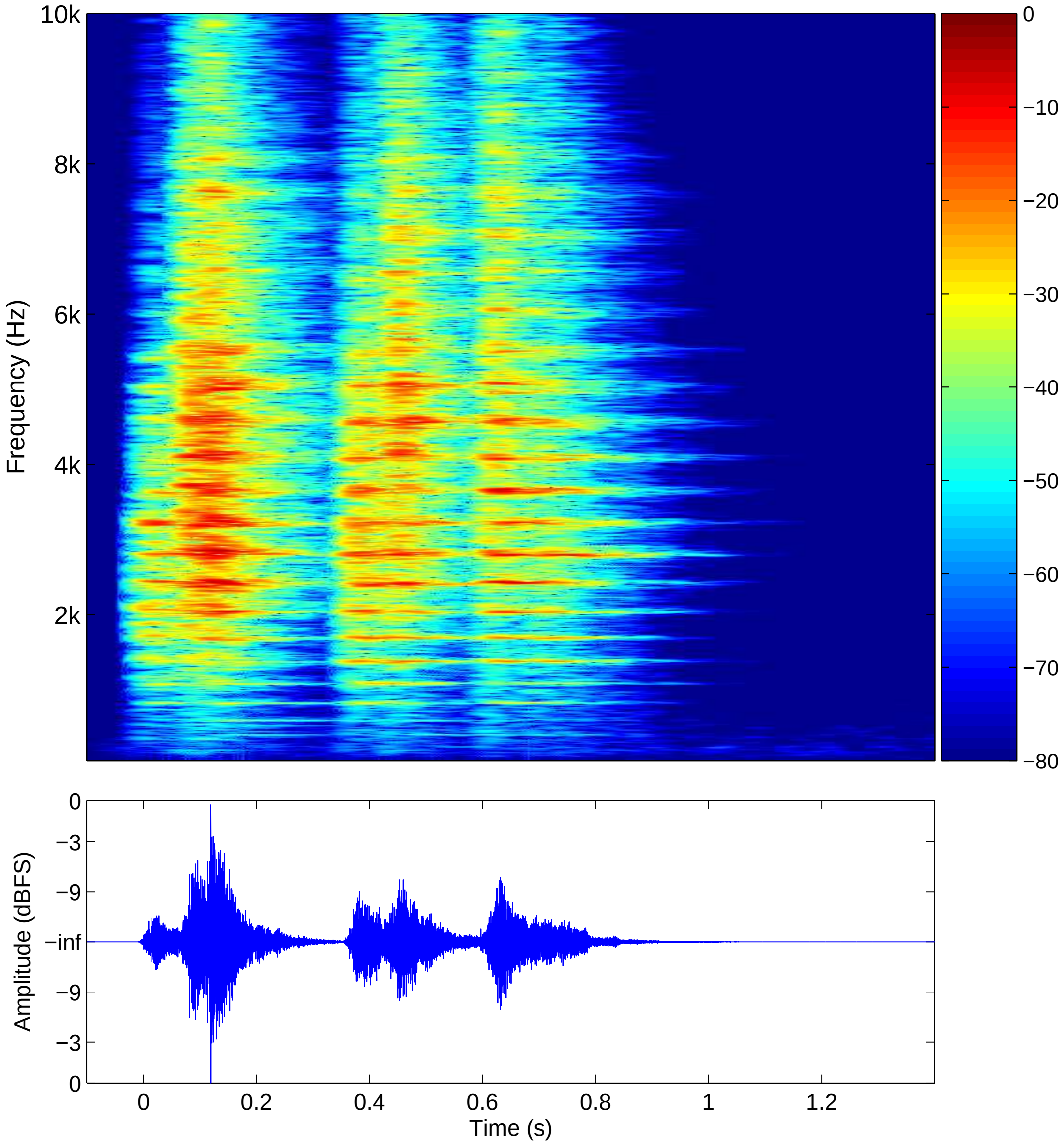


Figure B.8: Wooden dowel, 120 cm long, 13 mm diameter.

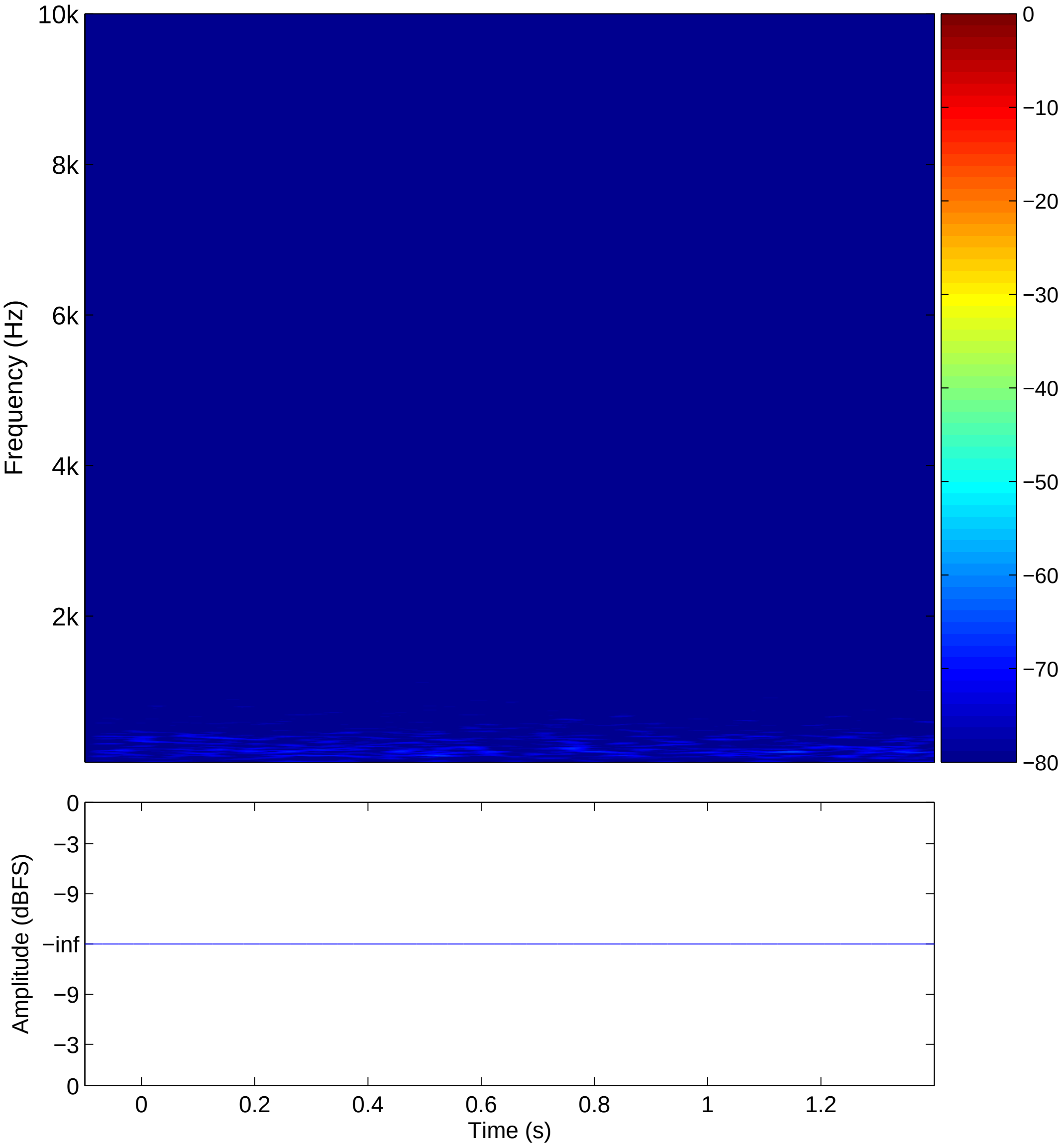


Figure B.9: Background noise.

Appendix C

Experiment Details: The Influence of Hearing Impairment and Hearing Aids on Impact Sound Perception

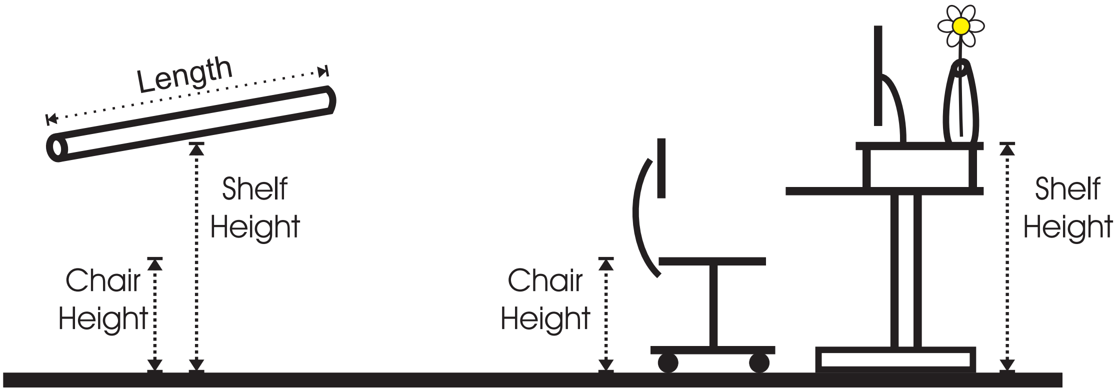
Listening Test Instructions

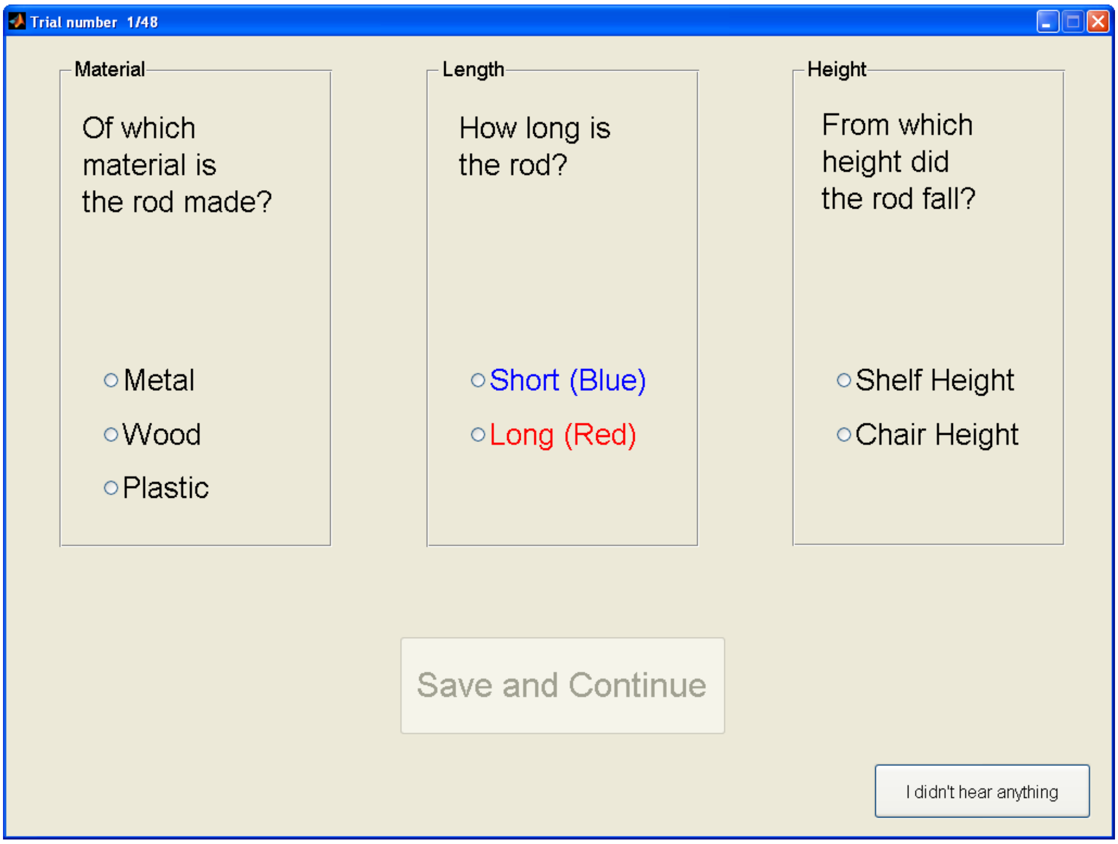
The listening test in which you are about to participate is part of a study used to help understand the hearing abilities of humans. In the experiment, you will hear various rods, both hollow and solid, dropped on the floor of the room in which you are sitting. Immediately after each rod is dropped, you will be asked to estimate the length of the rod, the height from which it was dropped, and the material of which it is made. These questions will be presented to you on the computer screen in front of you, and you should respond by using the mouse to click your choices. Attached to these instructions is an example of what the response form will look like. You may answer the questions in any order and should press the Save and Continue button when you are satisfied with your selections. Press the I didn't hear anything button if you have not hear a sound when the response menu appears.

The test will be split into four sessions, each lasting approximately fifteen minutes. In some of these sessions, you will be asked to wear hearing aids. In others, you will listen without hearing aids. Your task is identical in both cases. You will be allowed to take a break between each session for as long as you wish.

While you are wearing the hearing aids, please do not touch the volume control or any buttons on the hearing aids.

If you have any questions now or during the test, simply notify the test operator. You may feel free to stop the test at anytime.





Material

Of which material is the rod made?

○ Metal
○ Wood
○ Plastic

Length

How long is the rod?

○ Short (Blue)
○ Long (Red)

Height

From which height did the rod fall?

○ Shelf Height
○ Chair Height

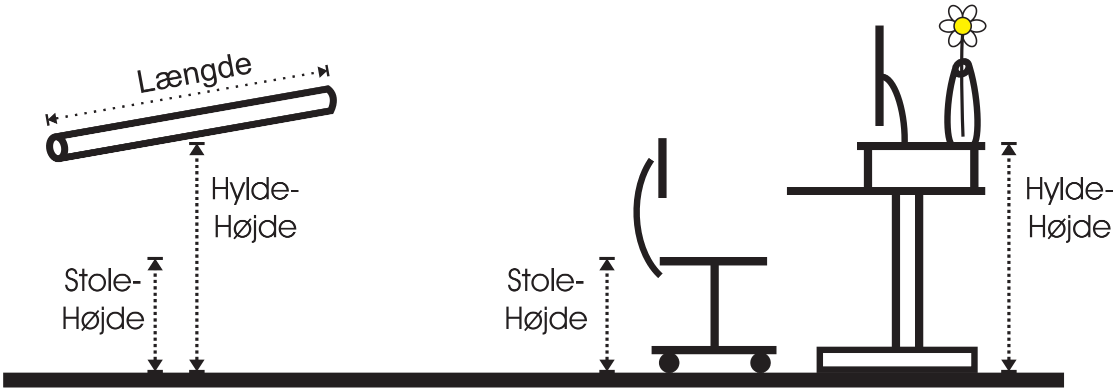
Vejledning til lyttetest (Danish Listening Test Instructions)

Lyttetesten som du skal deltage i er led i en større undersøgelse som skal hjælpe os til at forstå menneskets høresans. I undersøgelsen vil du høre forskellige stave, både hule og massive, som bliver tabt på gulvet i det rum hvor du sidder. Straks efter at staven er tabt på gulvet, vil du blive bedt om at skønne hvor lang du tror staven er, fra hvilken højde den blev tabt og hvilket materiale den er lavet af. Disse spørgsmål bliver stillet på computer skærmen, og du skal svare på spørgsmålene ved at bruge musen til at klikke på dine svar. Vedhæftet denne vejledning kan du se et eksempel på hvordan svarskærmen ser ud. Du kan svare på spørgsmålene i den rækkefølge du selv synes og når du er tilfreds med dine svar skal du trykke på Gem og Fortsæt . Tryk på Jeg har ikke hørt noget hvis du slet ikke har hørt noget når svarskærmen kommer frem.

Testen er delt op i fire dele, og hver del tager ca. 15 minutter. I visse dele af testen, bliver du bedt om at tage høreapparater på, og i andre dele af testen skal du høre uden, men din opgave er den samme. Der er mulighed for at holde pause mellem de fire dele af testen.

Når du har høreapparater på, må du ikke røre apparaterne - heller ikke for at justere lyden.

Hvis du har spørgsmål nu eller under testen, så spørg forsøgslederen. Du kan når som helst afbryde din deltagelse i undersøgelsen.



Hvad er staven lavet af?

○ Metal
○ Træ
○ Plastik

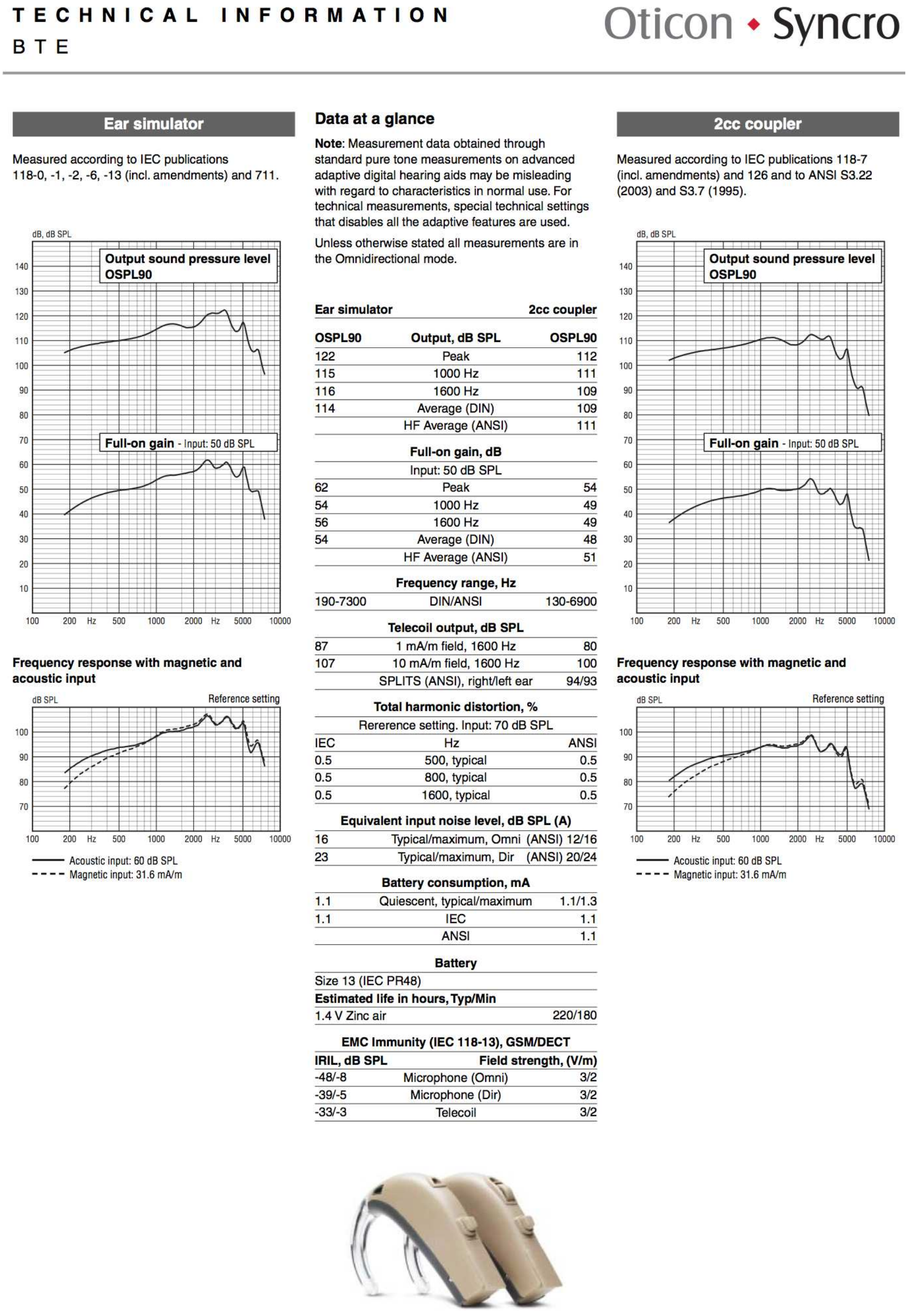
Hvor lang er staven?

○ Kort (Blå)
○ Lang (Rød)

Hvilken højde faldt staven fra?

○ Hylde-højde
○ Stole-højde

Hearing Aid for Normal-Hearing Subjects



TECHNICAL INFORMATION

BTE

Oticon ◆ Syncro

Ear simulator

Measured according to IEC publications 118-0, -1, -2, -6, -13 (incl. amendments) and 711.

Data at a glance

Note: Measurement data obtained through standard pure tone measurements on advanced adaptive digital hearing aids may be misleading with regard to characteristics in normal use. For technical measurements, special technical settings that disables all the adaptive features are used.

Unless otherwise stated all measurements are in the Omnidirectional mode.

2cc coupler

Measured according to IEC publications 118-7 (incl. amendments) and 126 and to ANSI S3.22 (2003) and S3.7 (1995).

Ear simulator		**2cc coupler**
OSPL90	**Output, dB SPL**	**OSPL90**
122	Peak	112
115	1000 Hz	111
116	1600 Hz	109
114	Average (DIN)	109
	HF Average (ANSI)	111
	Full-on gain, dB	
	Input: 50 dB SPL	
62	Peak	54
54	1000 Hz	49
56	1600 Hz	49
54	Average (DIN)	48
	HF Average (ANSI)	51
	Frequency range, Hz	
190-7300	DIN/ANSI	130-6900
	Telecoil output, dB SPL	
87	1 mA/m field, 1600 Hz	80
107	10 mA/m field, 1600 Hz	100
	SPLITS (ANSI), right/left ear	94/93
	Total harmonic distortion, %	
	Rererence setting. Input: 70 dB SPL	
IEC	Hz	ANSI
0.5	500, typical	0.5
0.5	800, typical	0.5
0.5	1600, typical	0.5
	Equivalent input noise level, dB SPL (A)	
16	Typical/maximum, Omni	(ANSI) 12/16
23	Typical/maximum, Dir	(ANSI) 20/24
	Battery consumption, mA	
1.1	Quiescent, typical/maximum	1.1/1.3
1.1	IEC	1.1
	ANSI	1.1
	Battery	
Size 13 (IEC PR48)		
Estimated life in hours, Typ/Min		
1.4 V Zinc air		220/180
	EMC Immunity (IEC 118-13), GSM/DECT	
IRIL, dB SPL		**Field strength, (V/m)**
-48/-8	Microphone (Omni)	3/2
-39/-5	Microphone (Dir)	3/2
-33/-3	Telecoil	3/2

Stimuli

Figures C.1 through C.24 show spectrograms and signal amplitude versus time plots of 24 bit recordings made from typical stimuli in the listening test described in Chapter 5. The signals are from the right ear of a Brüel & Kjær acoustic manikin used to make the recordings. They represent the signals at the position of the right eardrum. The stimuli are shown in the figures without the added fan noise that was present during the actual tests, but a spectrogram and plot of the amplitude versus time signal for the fan noise is shown in Figure C.25. Figure C.26 shows the typical background noise level without the fan. The plots in Figures C.1 through C.26 have been scaled so that the spectrogram colors can be compared between figures. The maximum decibel levels (relative) in the spectrograms are shown in red and correspond to 0 dBFS. The minimum values, which correspond to -105 dBFS, are shown in dark blue.

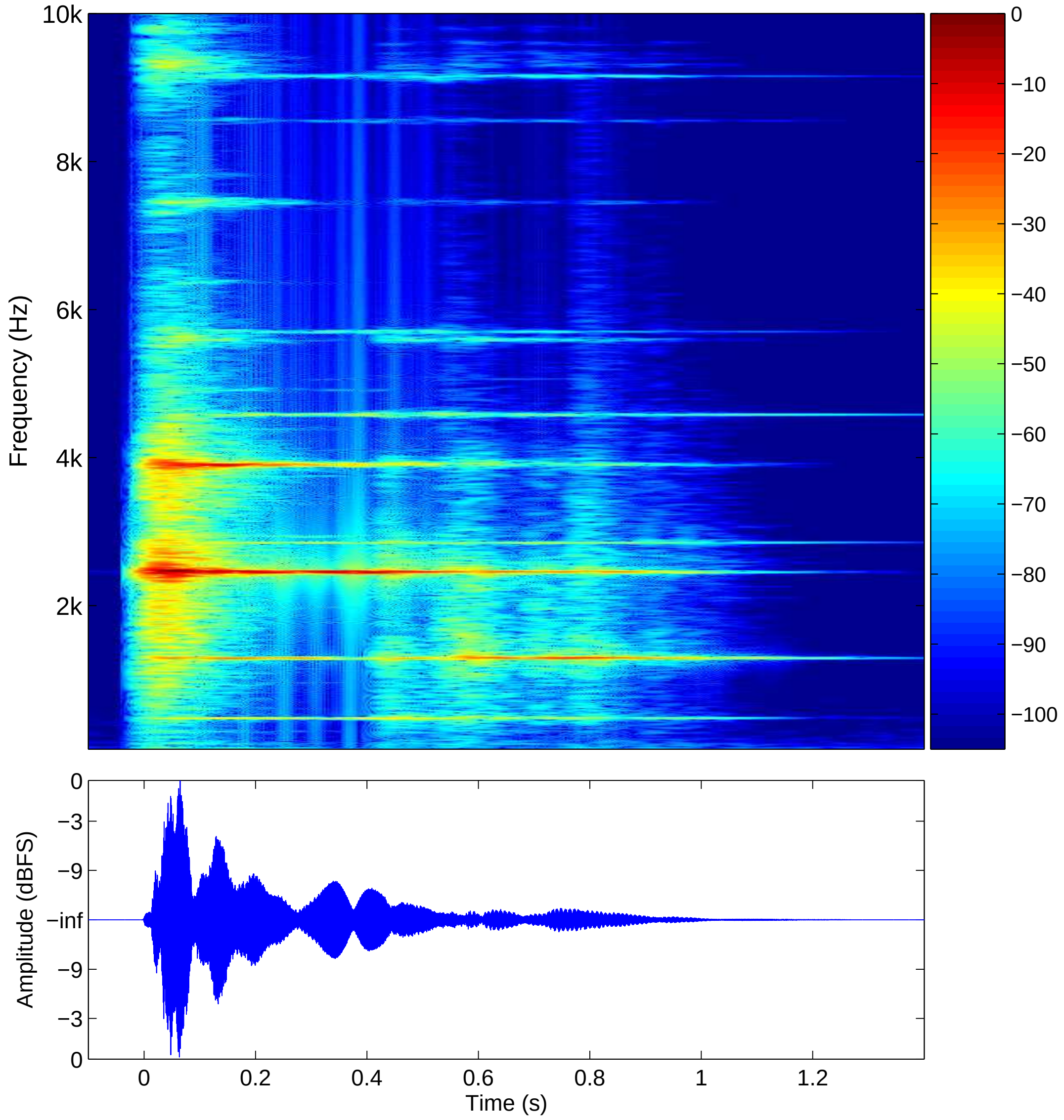


Figure C.1: Metal, thick, long (55 cm), shelf height.

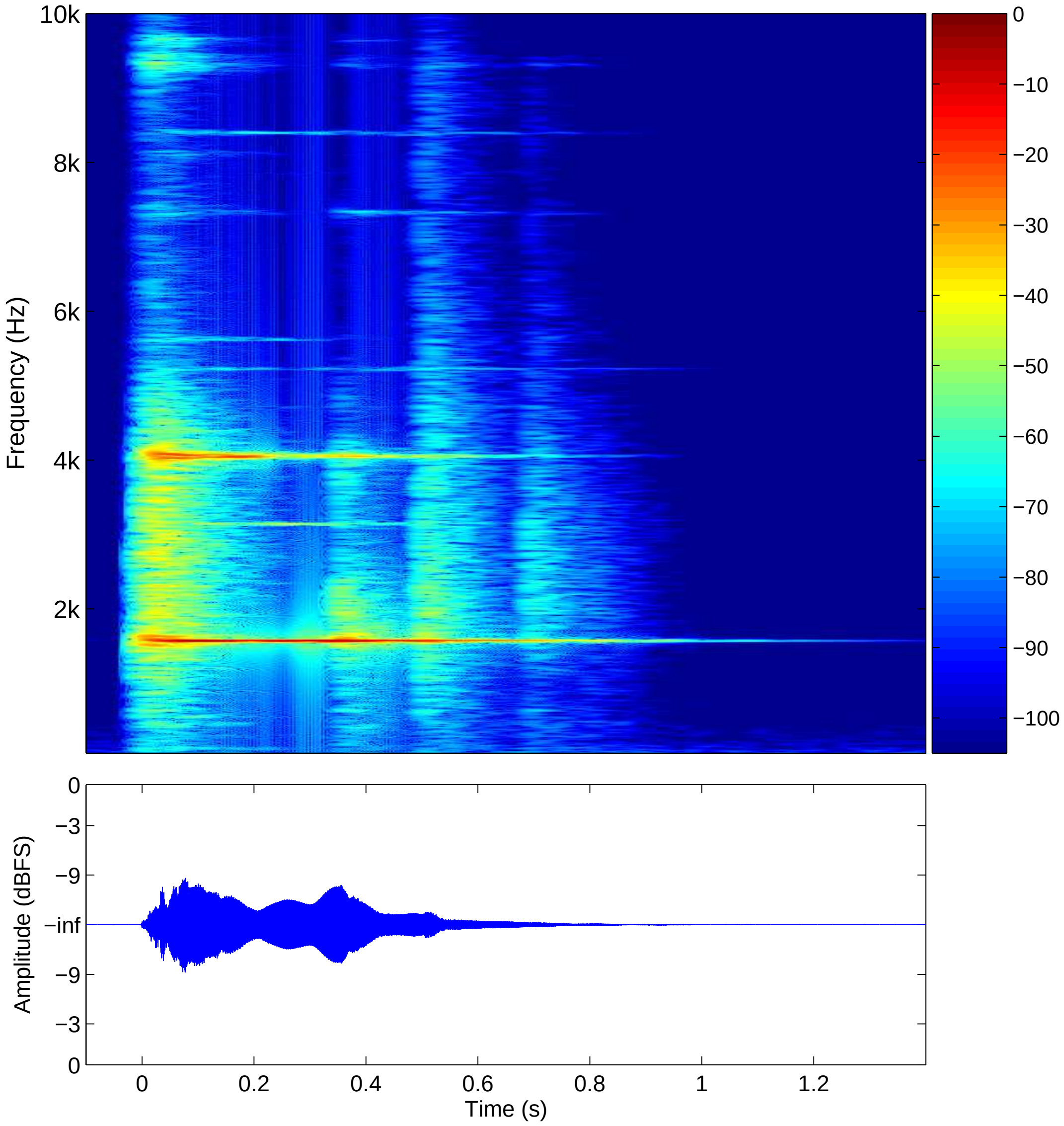


Figure C.2: Metal, thick, short (30 cm), shelf height.

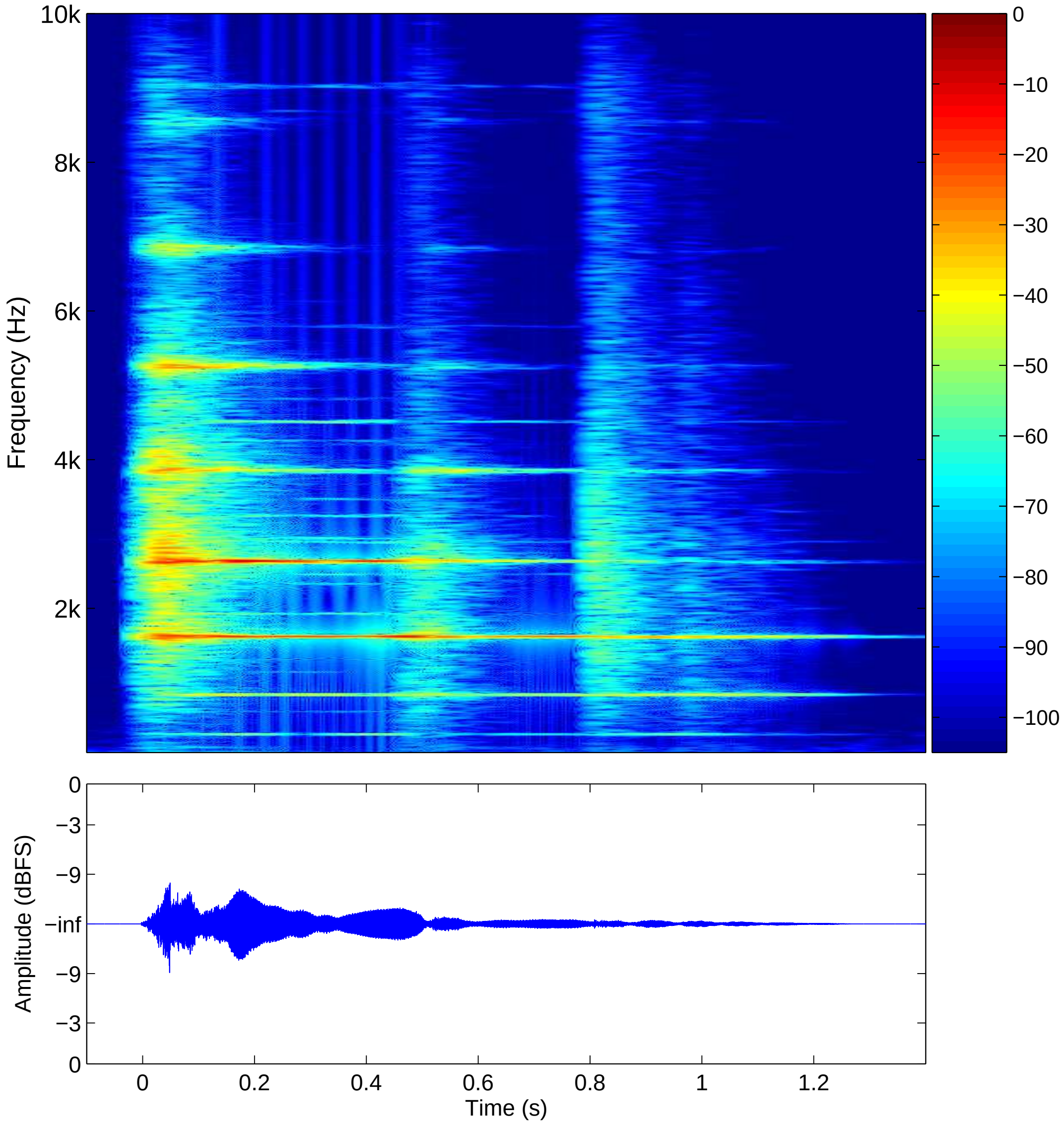


Figure C.3: Metal, thin, long (55 cm), shelf height.

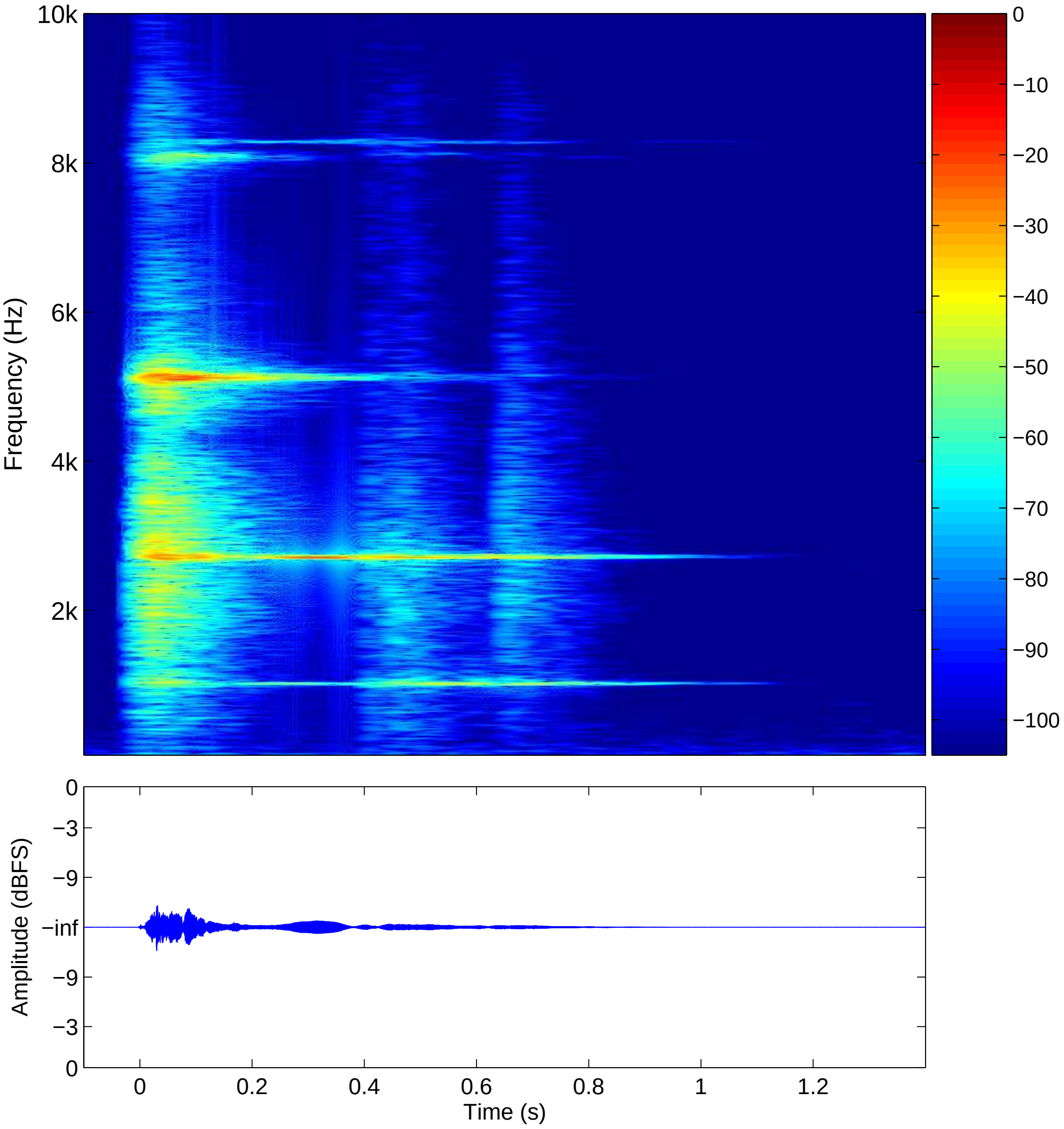


Figure C.4: Metal, thin, short (30 cm), shelf height.

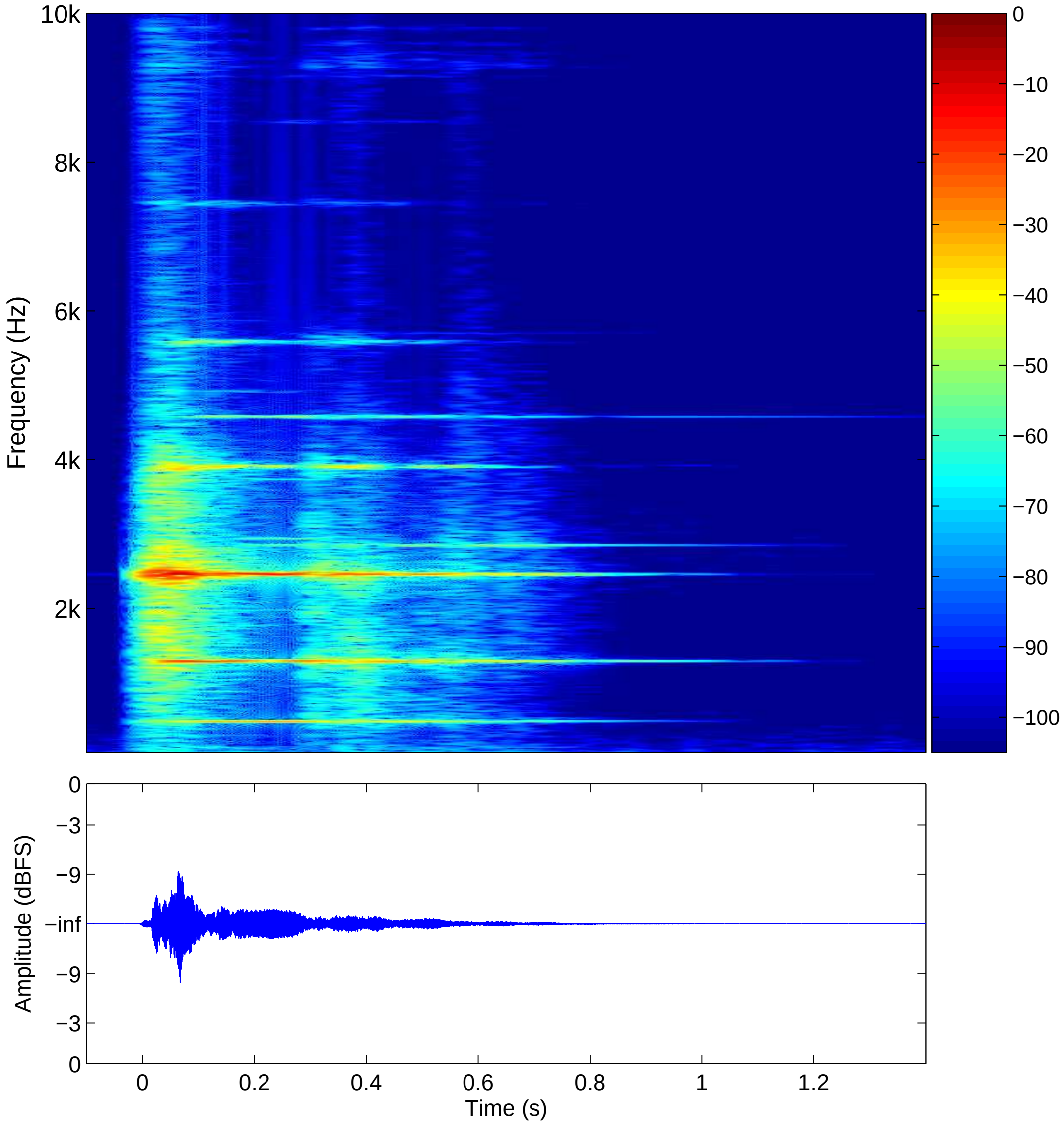


Figure C.5: Metal, thick, long (55 cm), chair height.

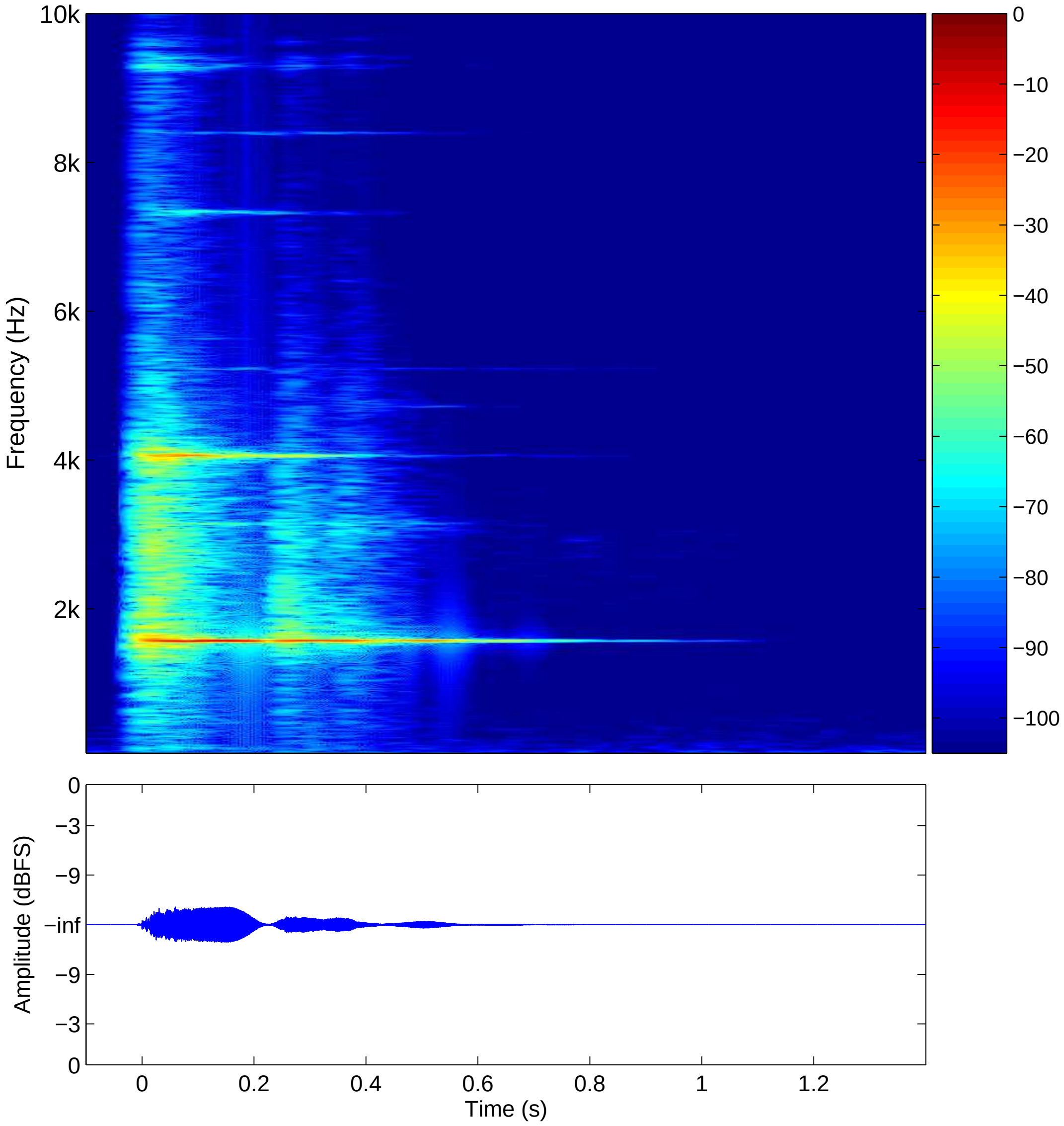


Figure C.6: Metal, thick, short (30 cm), chair height.

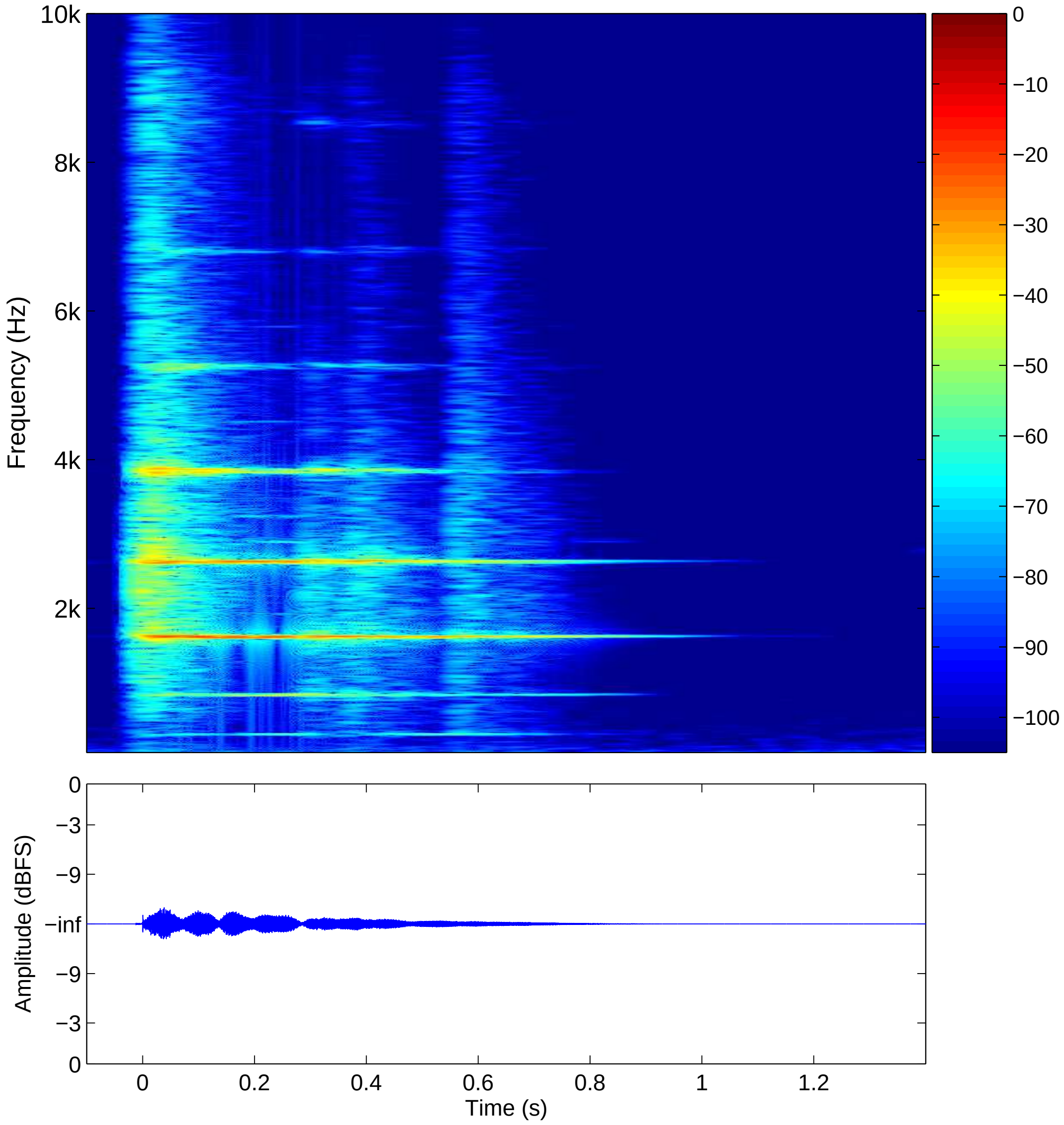


Figure C.7: Metal, thin, long (55 cm), chair height.

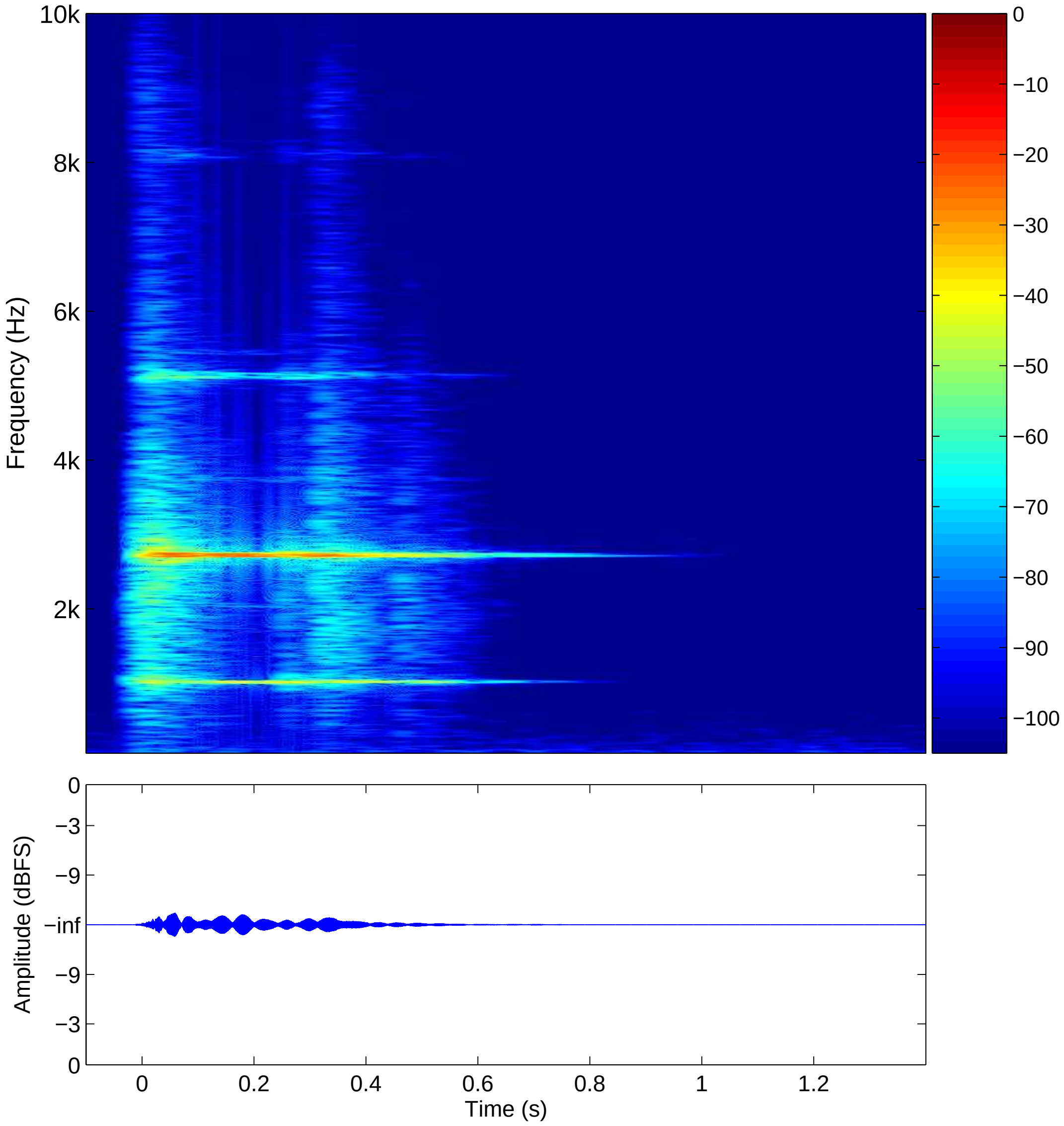


Figure C.8: Metal, thin, short (30 cm), chair height.

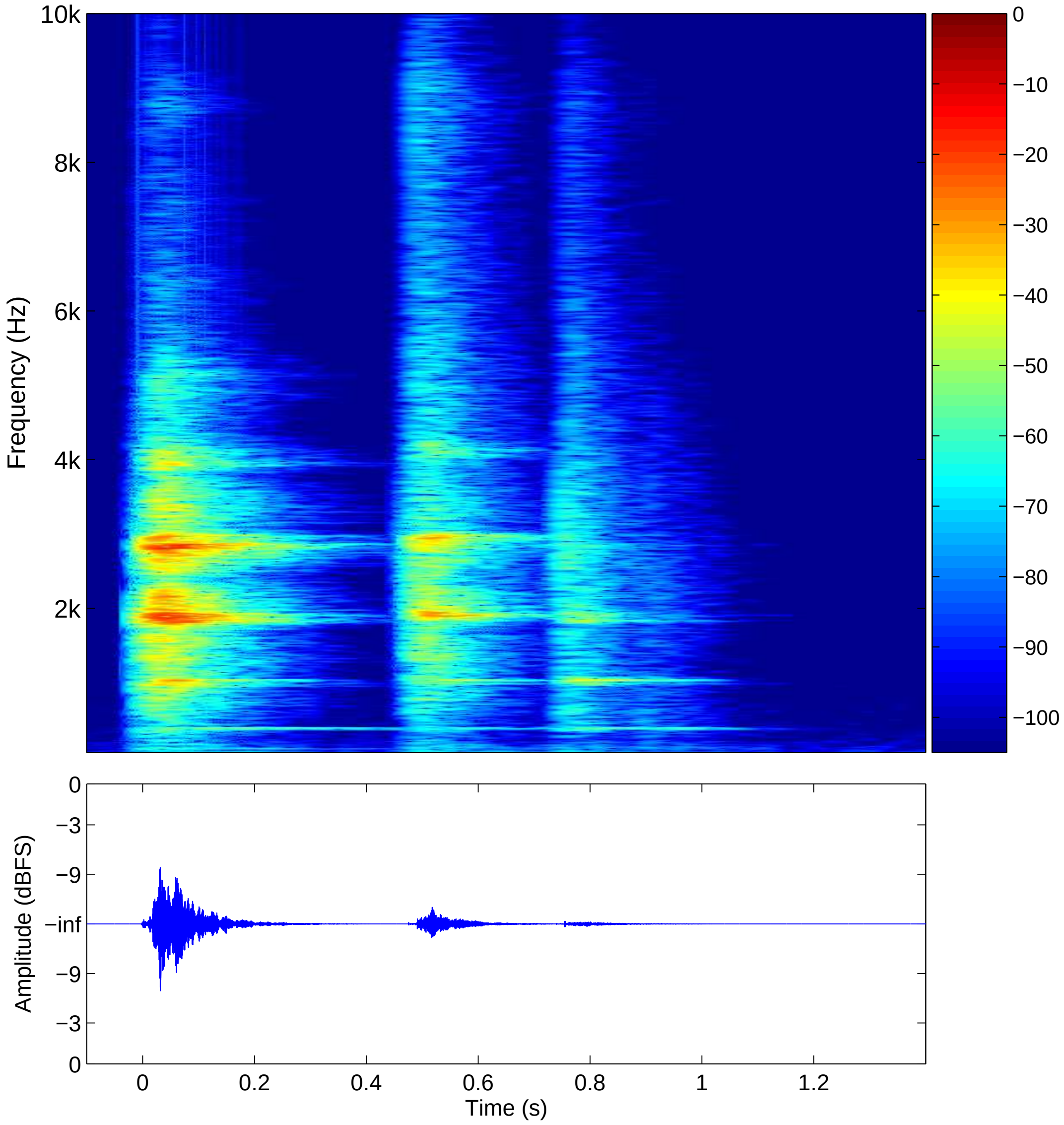


Figure C.9: Wood, thick, long (55 cm), shelf height.

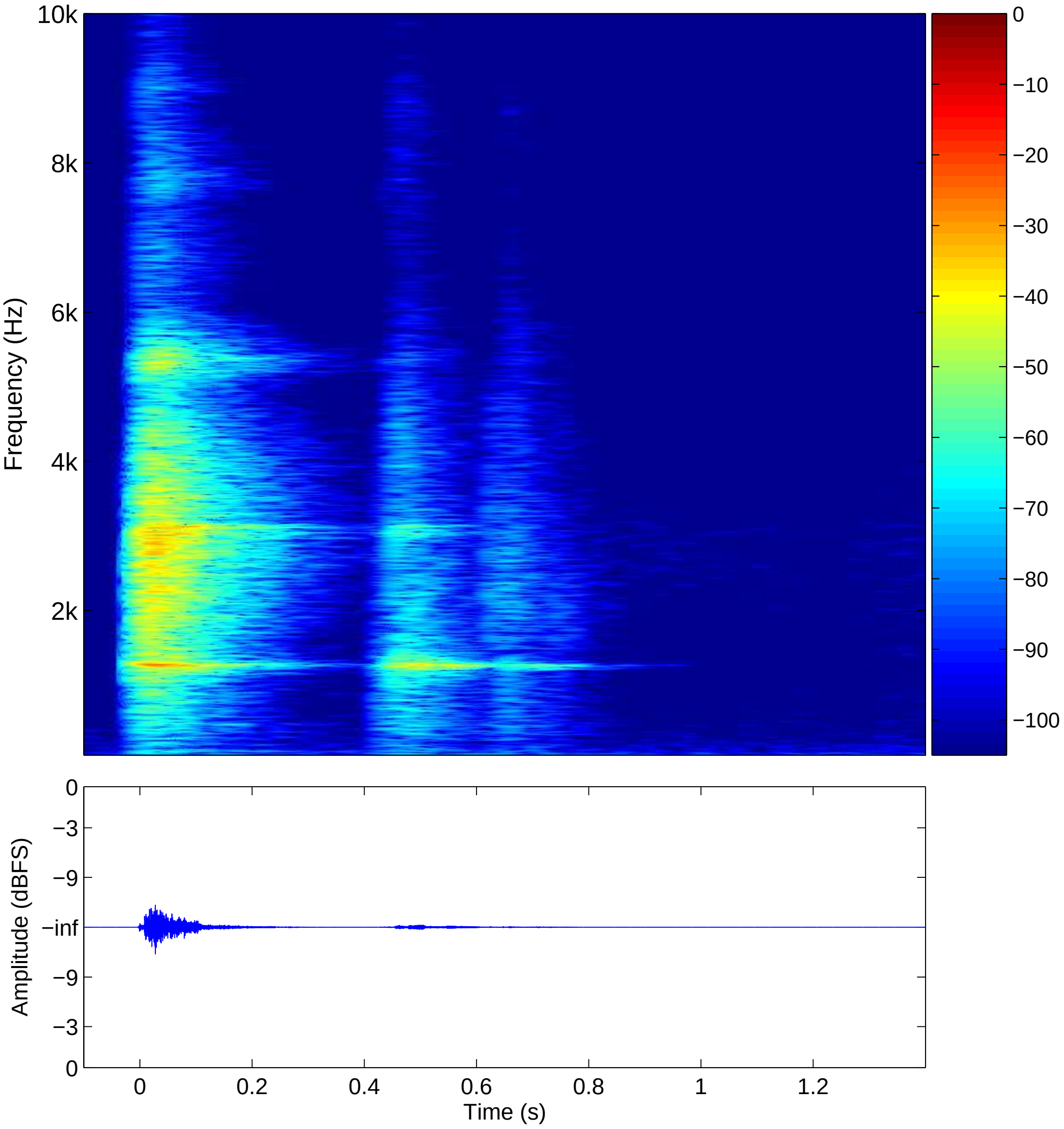


Figure C.10: Wood, thick, short (30 cm), shelf height.

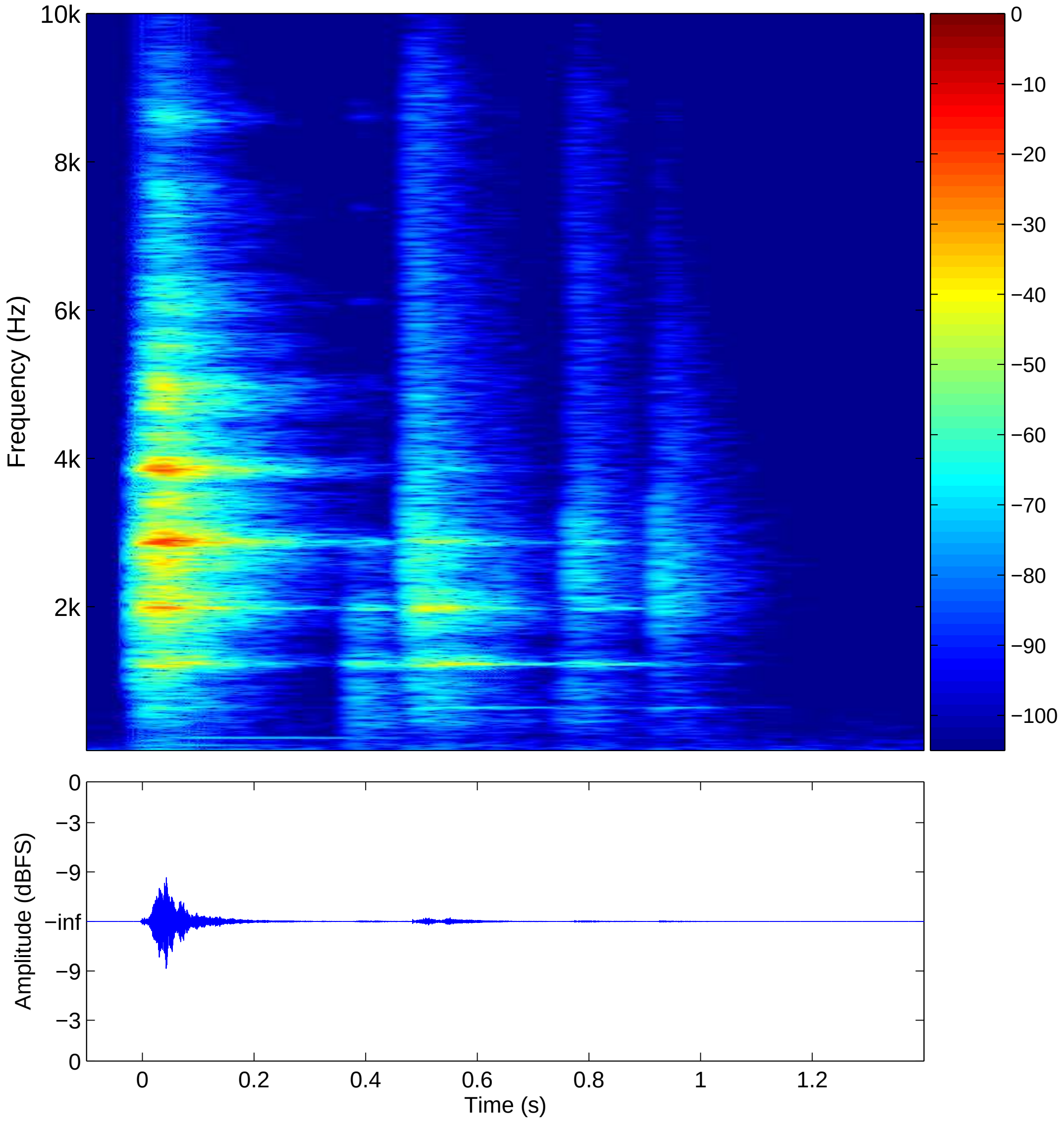


Figure C.11: Wood, thin, long (55 cm), shelf height.

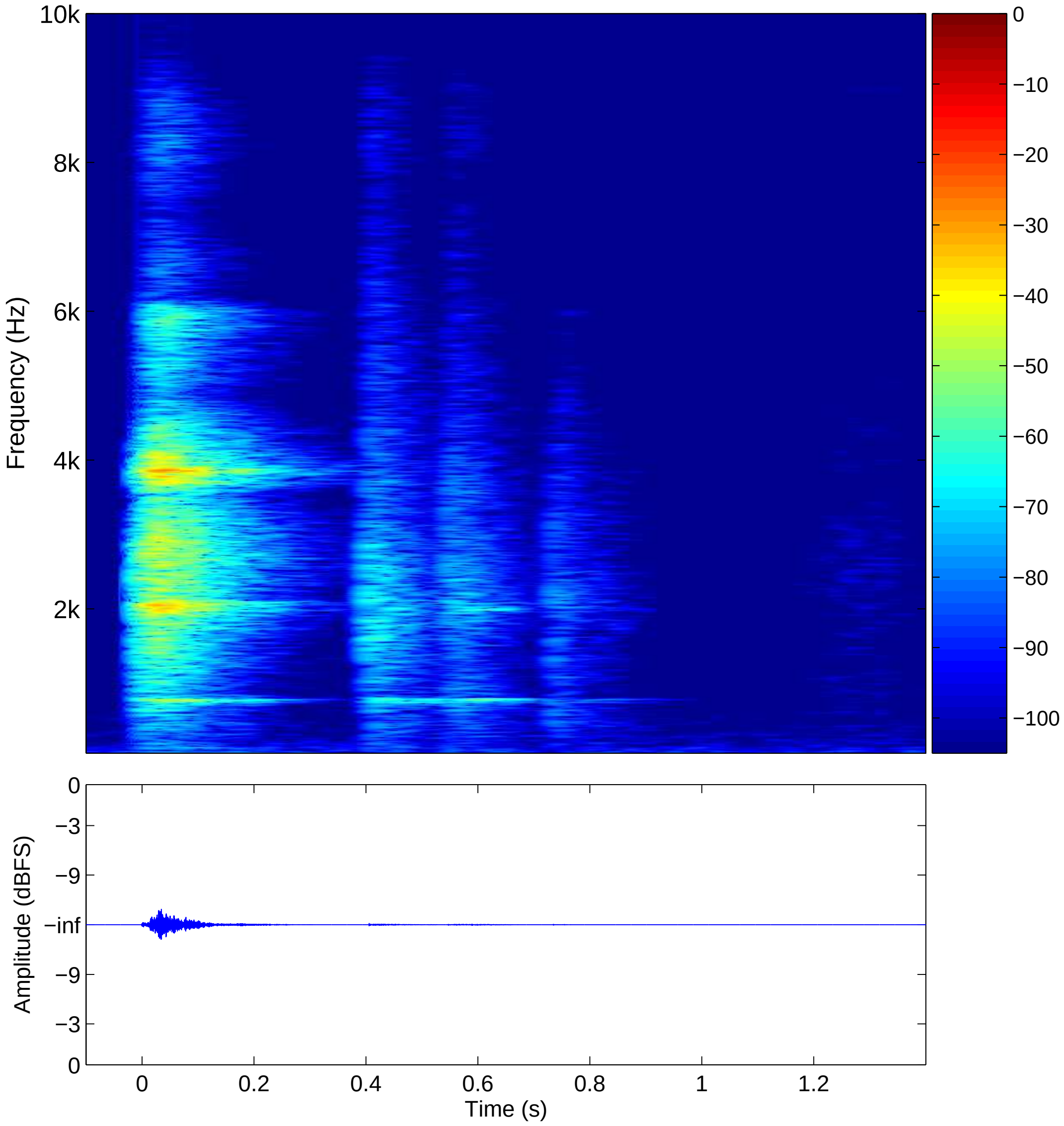


Figure C.12: Wood, thin, short (30 cm), shelf height.

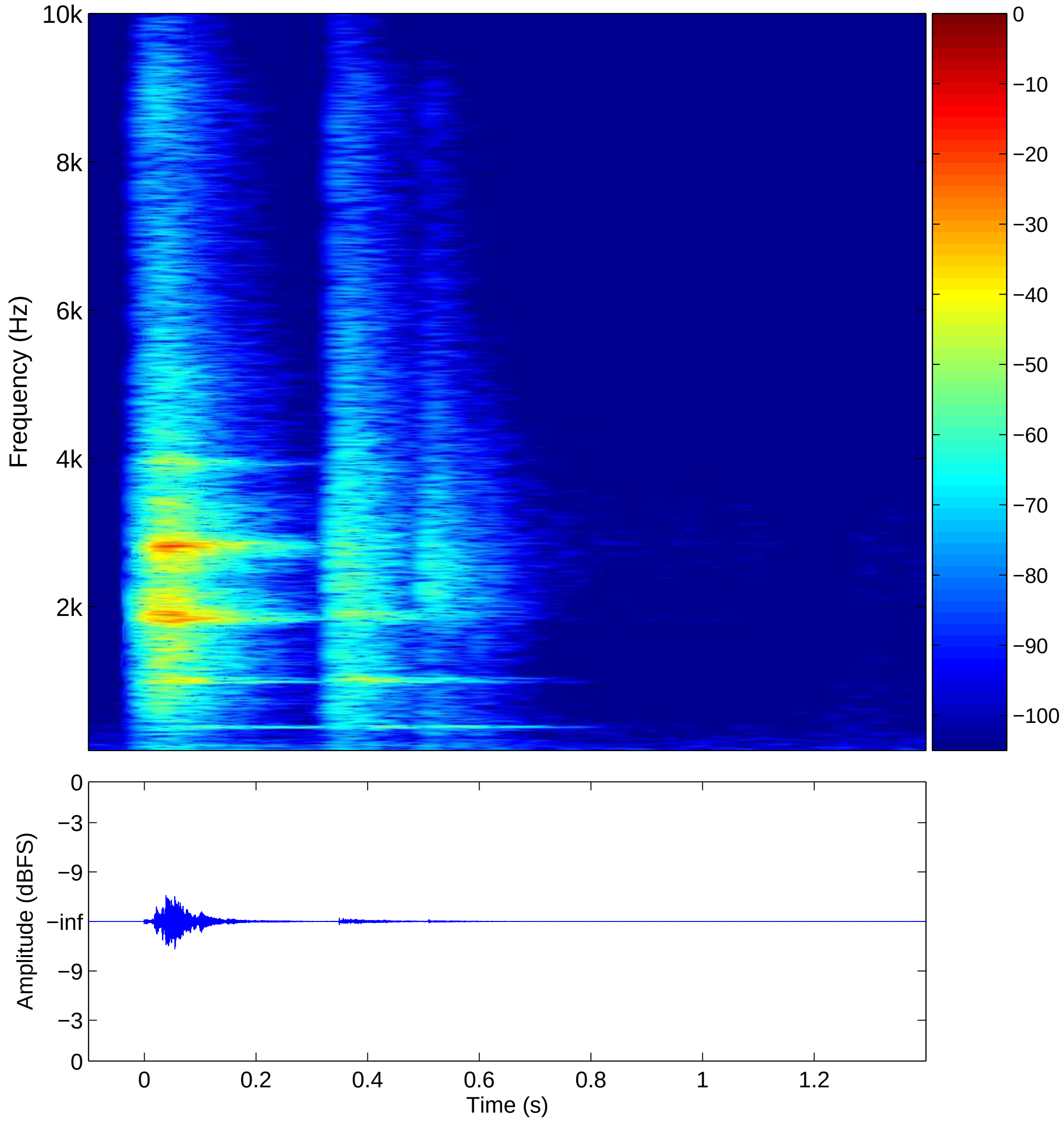


Figure C.13: Wood, thick, long (55 cm), chair height.

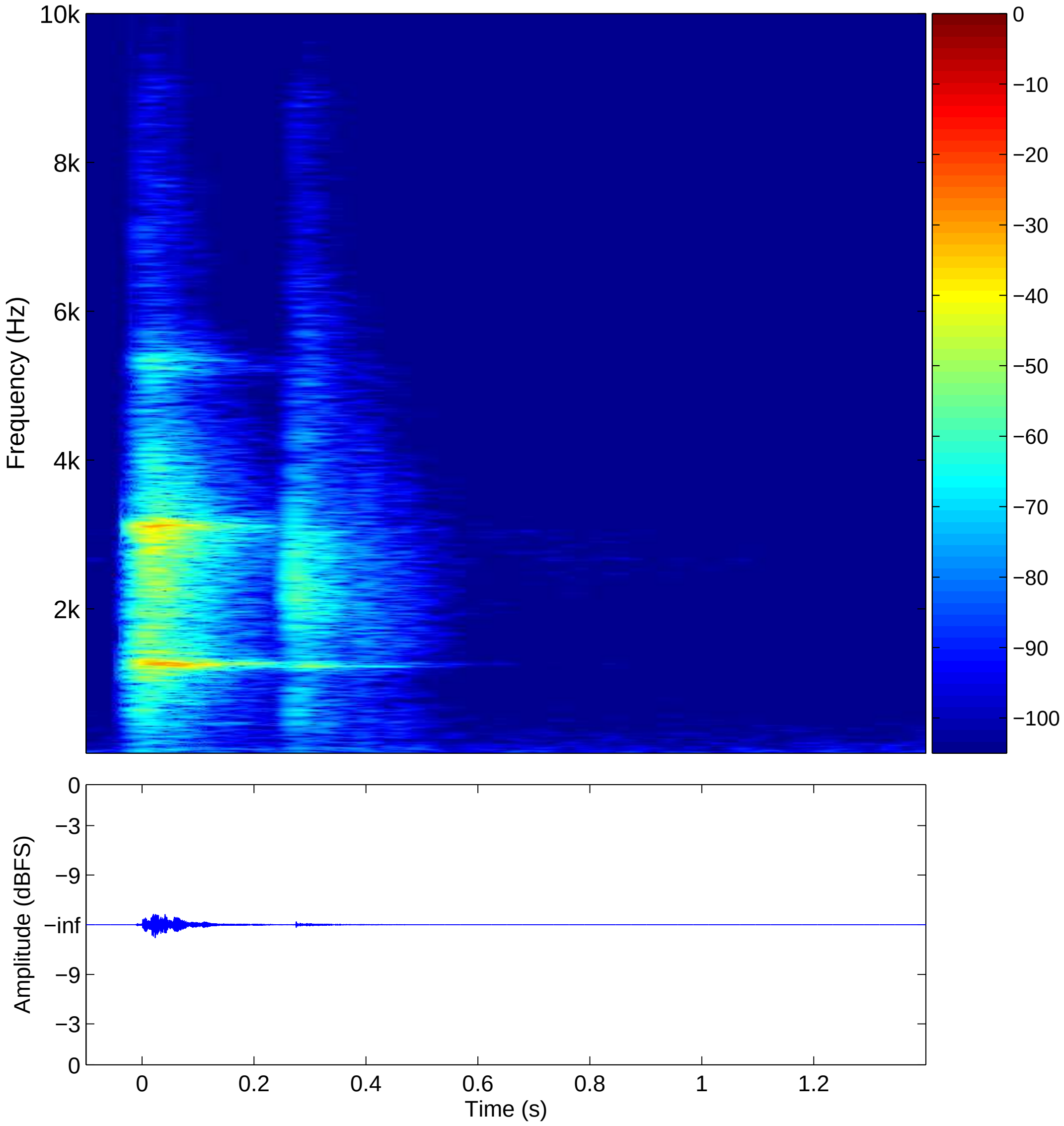


Figure C.14: Wood, thick, short (30 cm), chair height.

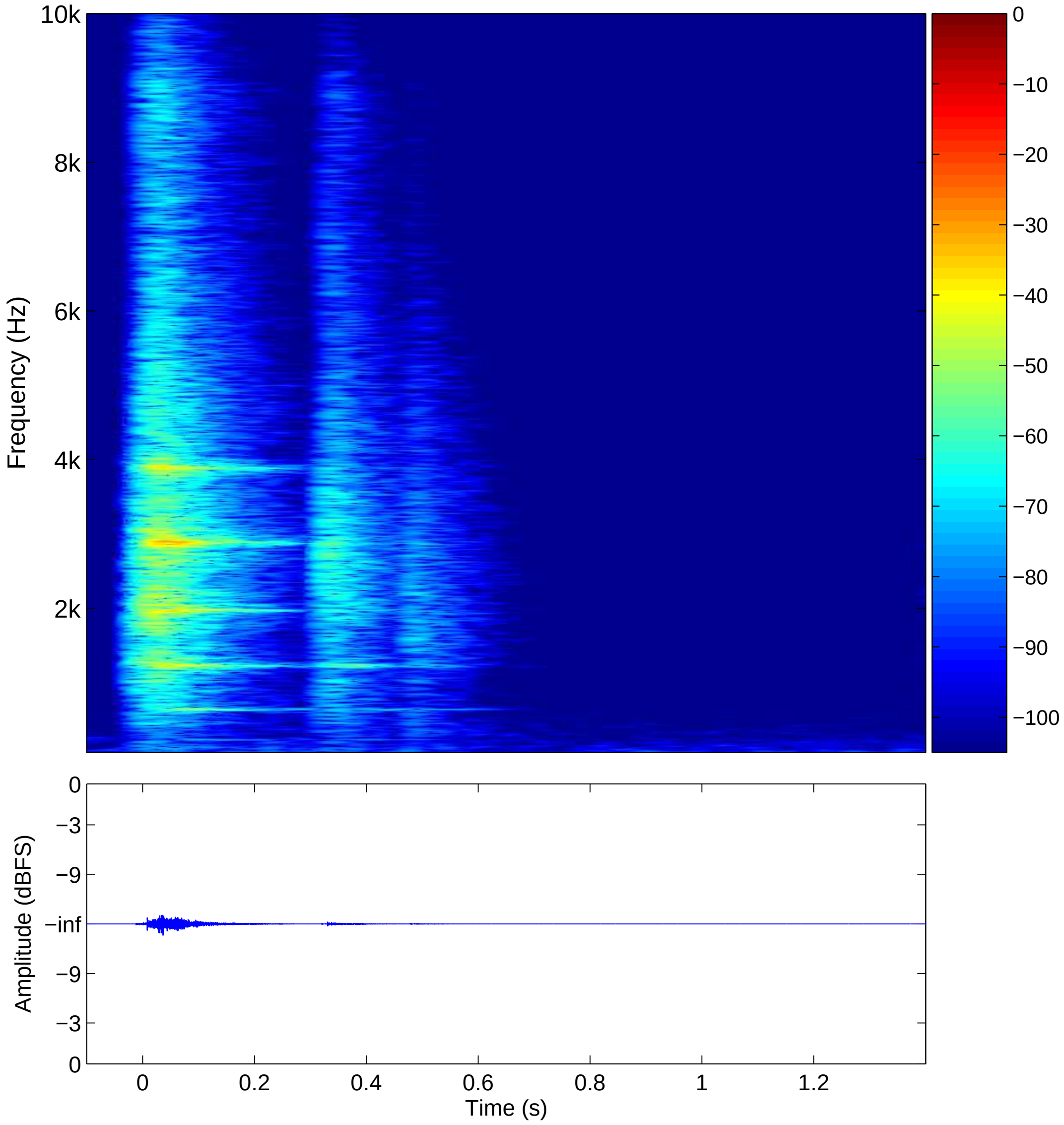


Figure C.15: Wood, thin, long (55 cm), chair height.

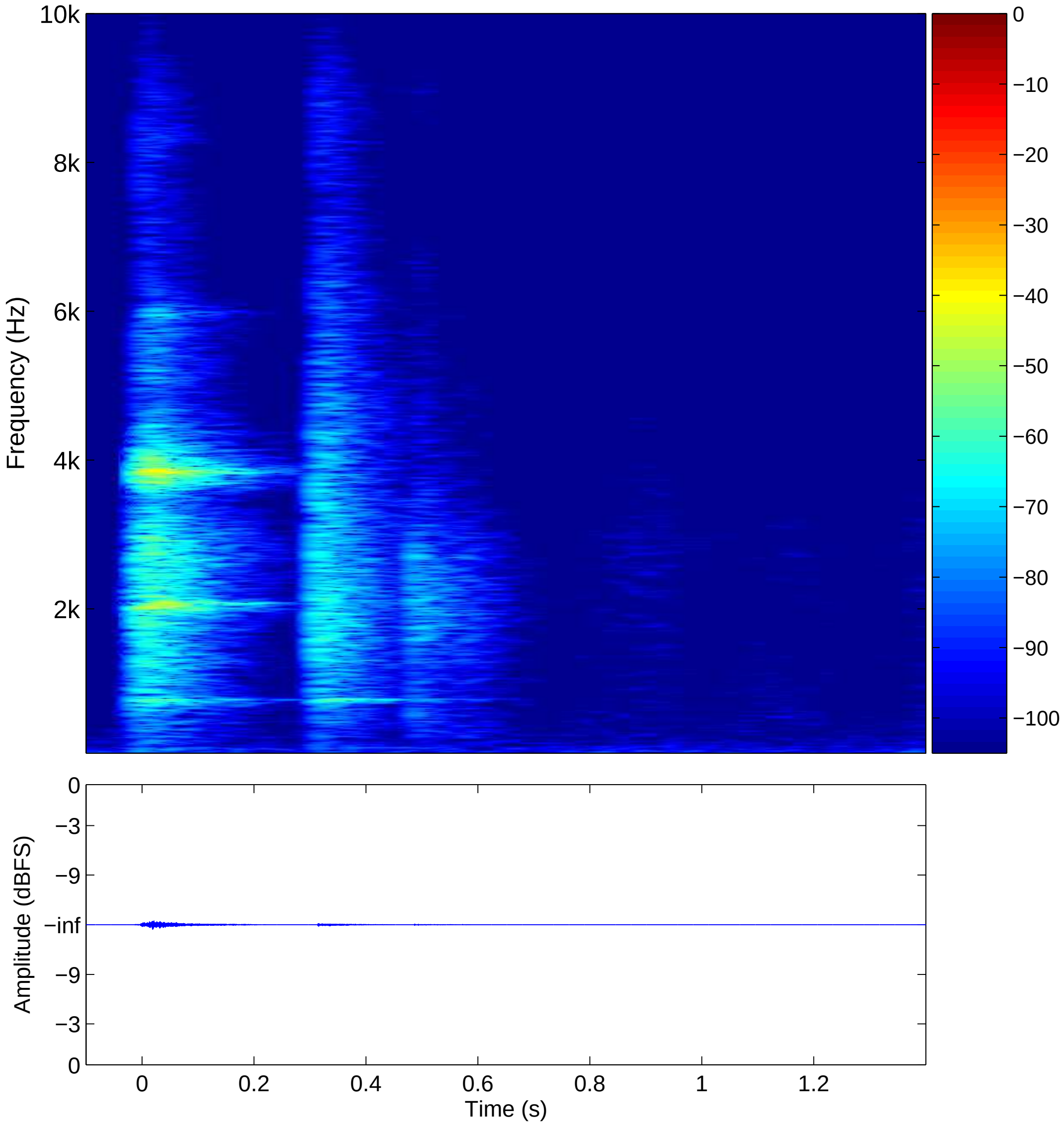


Figure C.16: Wood, thin, short (30 cm), chair height.

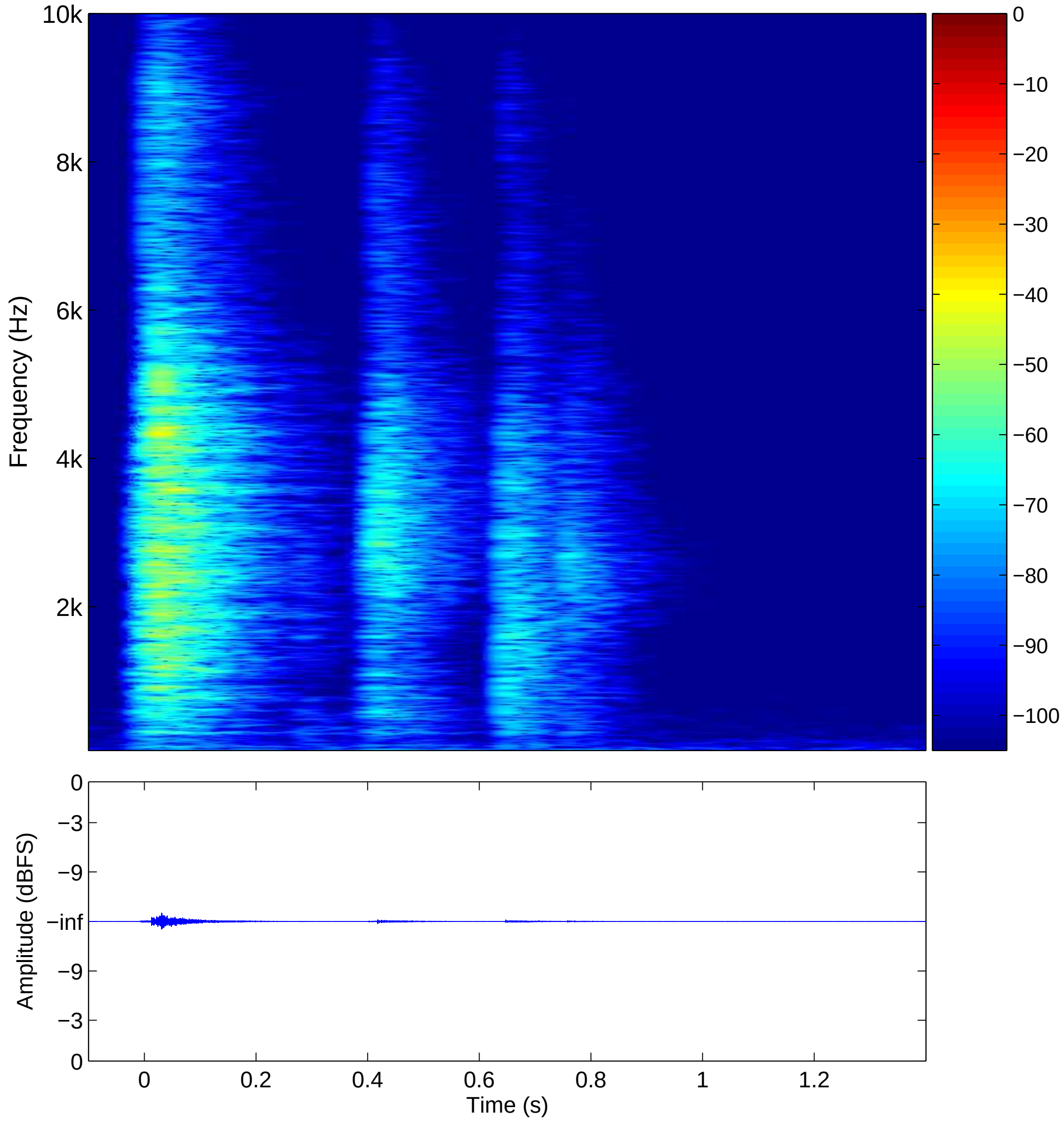


Figure C.17: Plastic, thick, long (55 cm), shelf height.

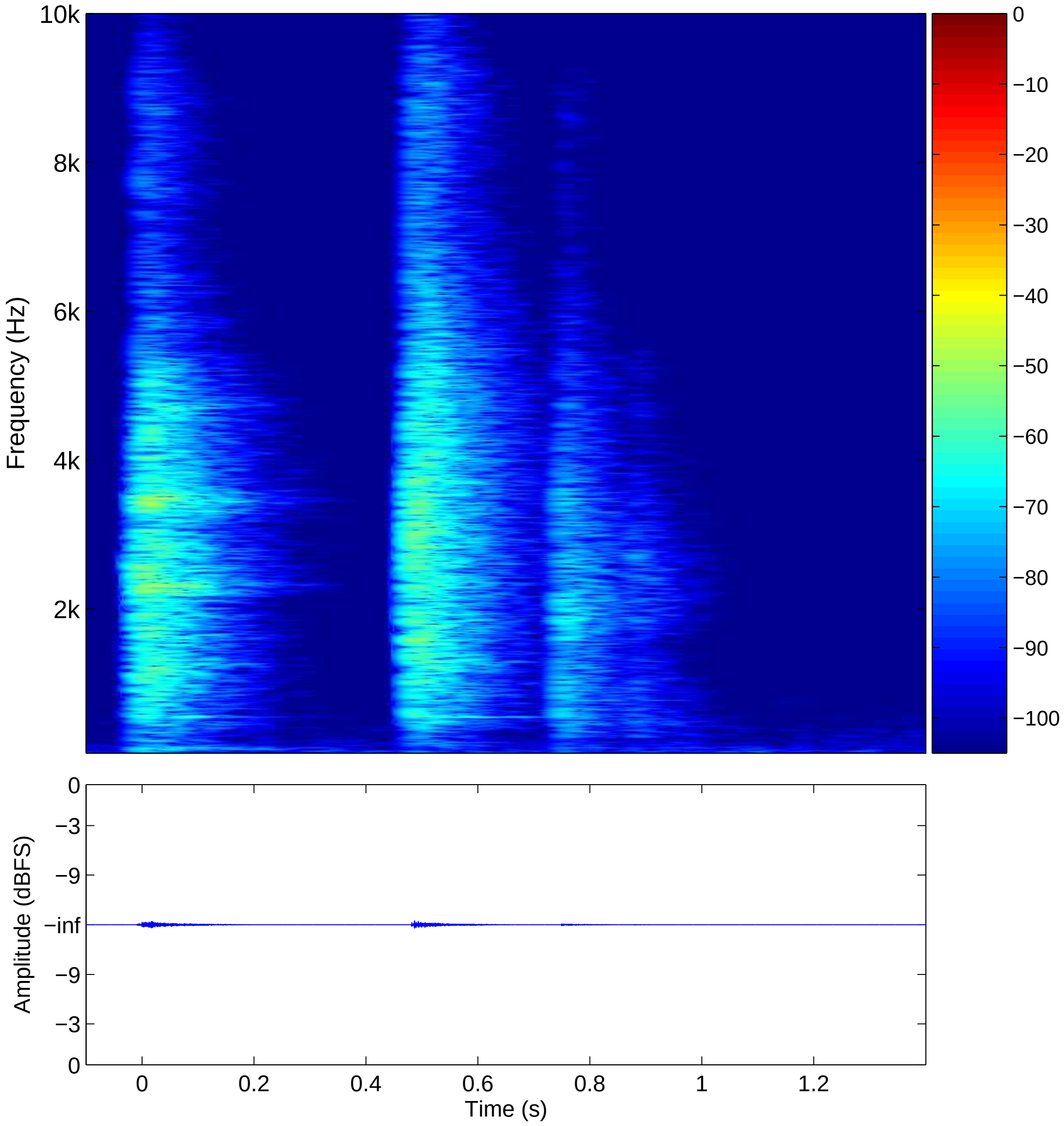


Figure C.18: Plastic, thick, short (30 cm), shelf height.

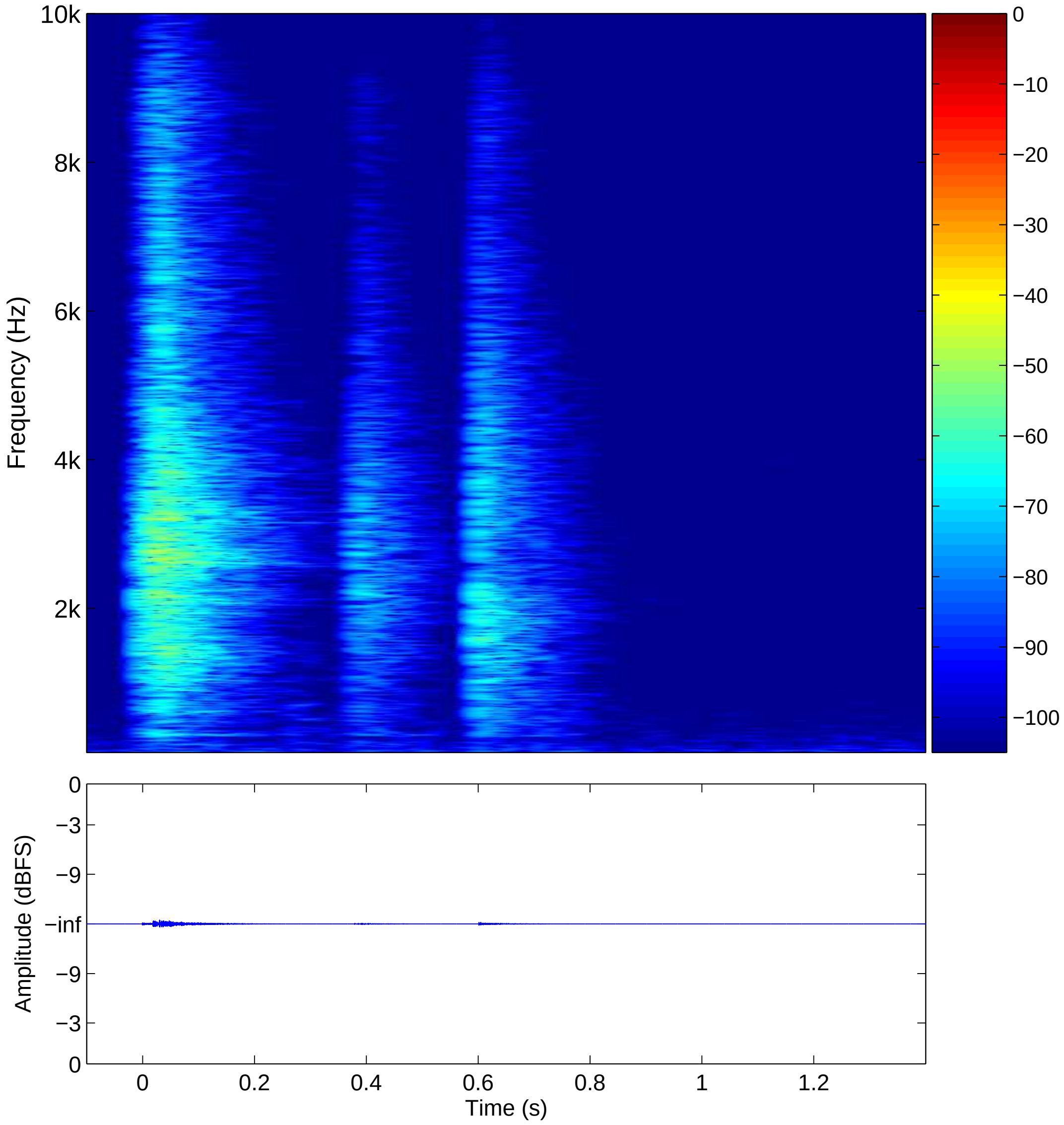


Figure C.19: Plastic, thin, long (55 cm), shelf height.

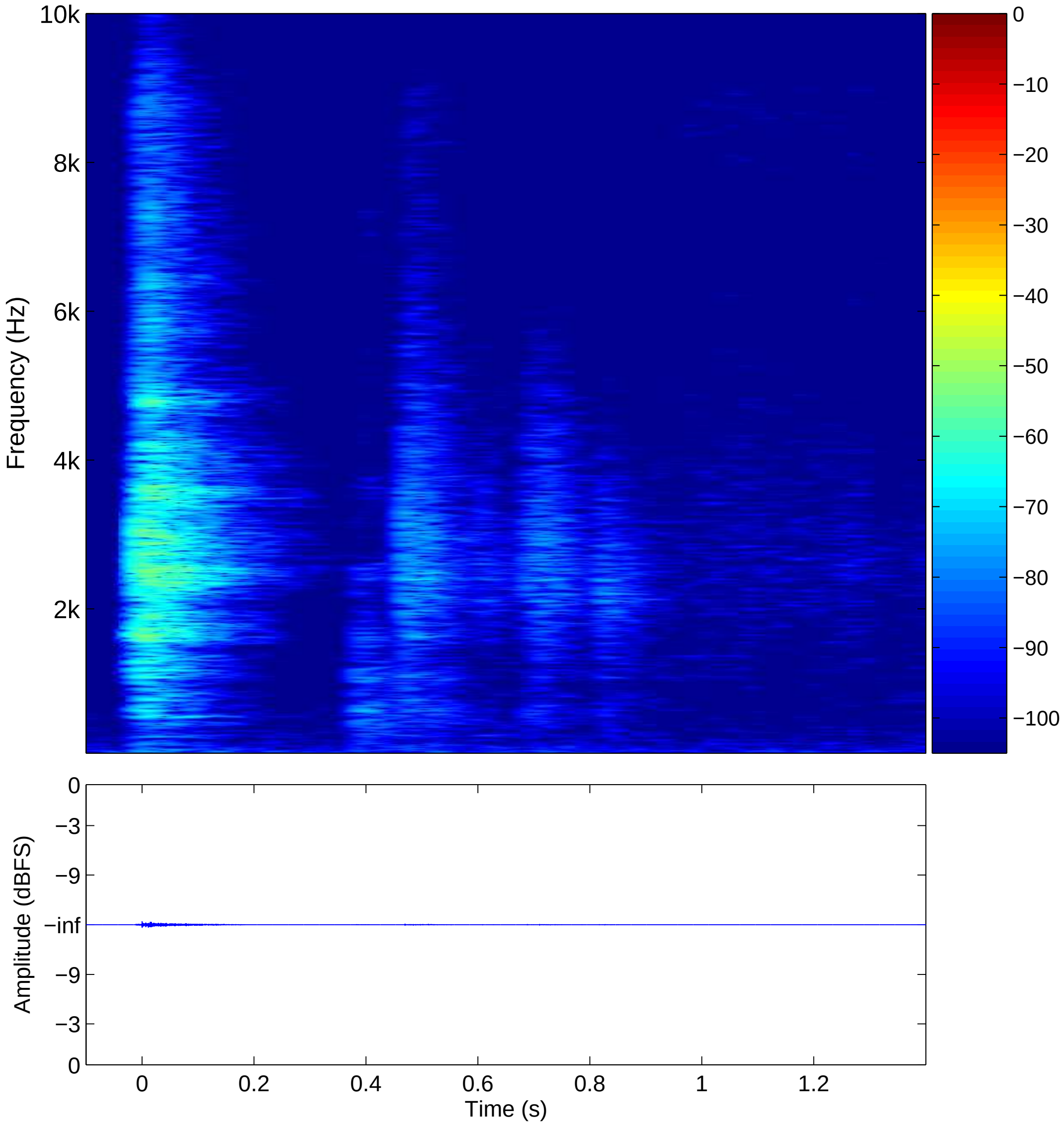


Figure C.20: Plastic, thin, short (30 cm), shelf height.

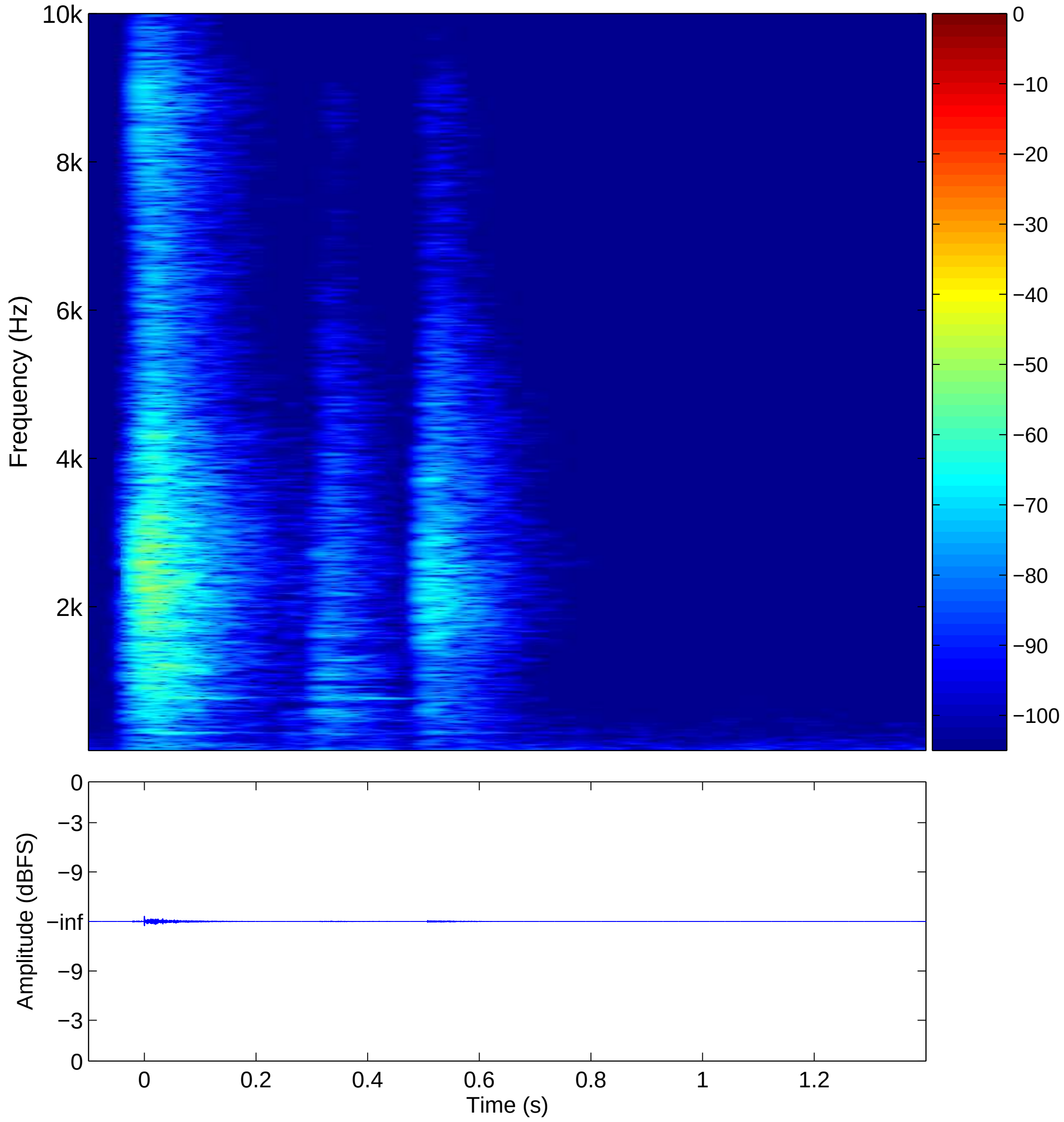


Figure C.21: Plastic, thick, long (55 cm), chair height.

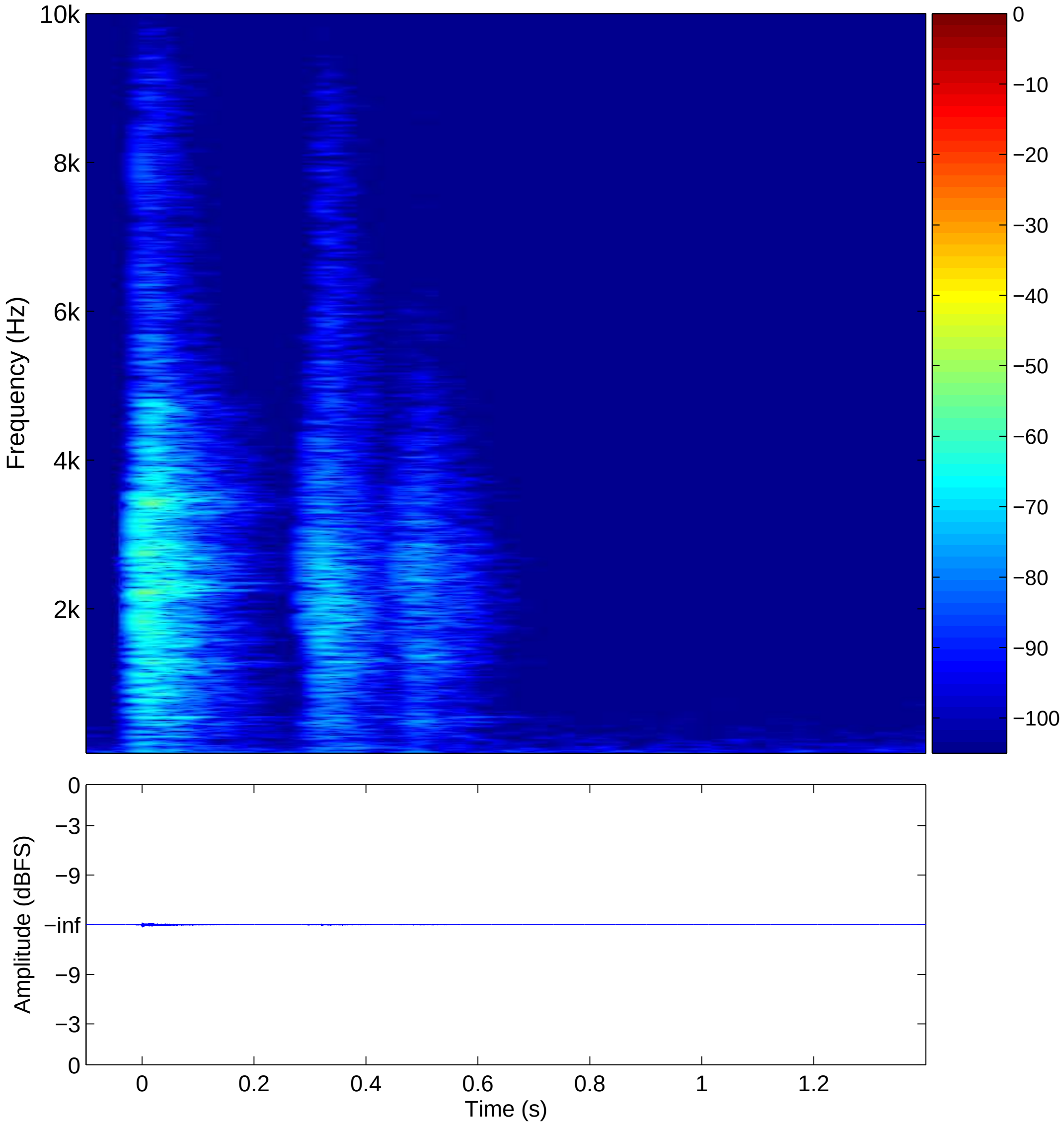


Figure C.22: Plastic, thick, short (30 cm), chair height.

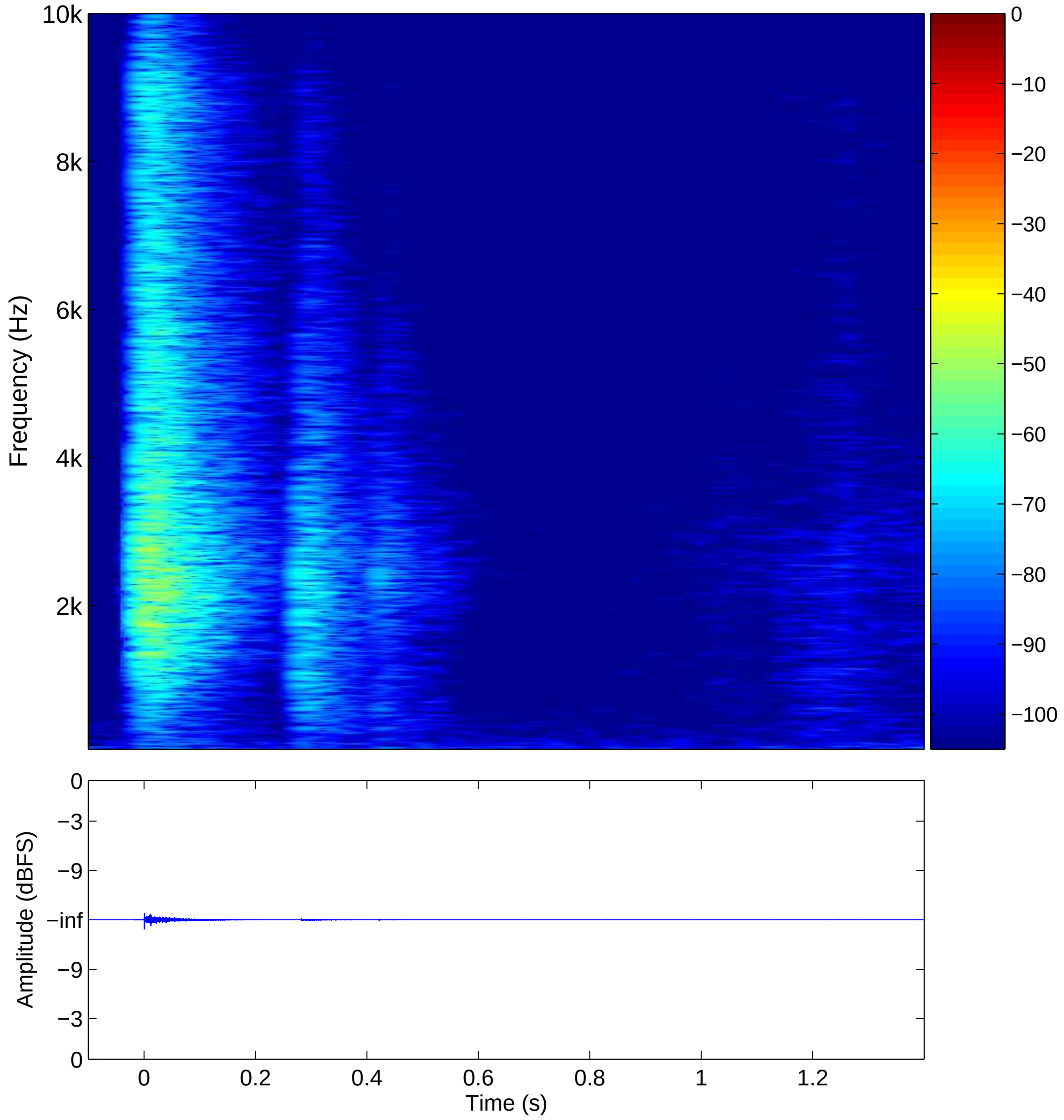


Figure C.23: Plastic, thin, long (55 cm), chair height.

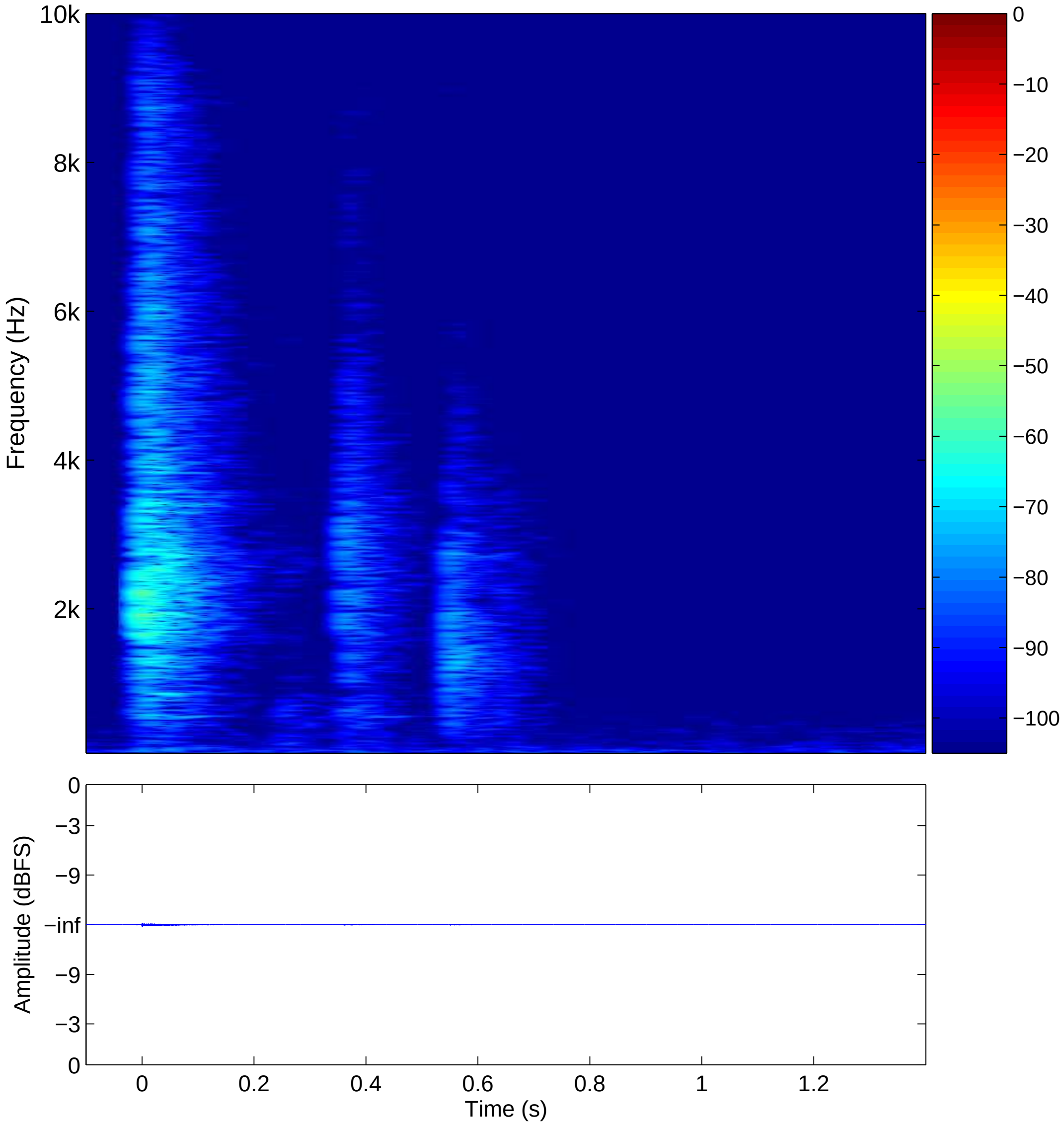


Figure C.24: Plastic, thin, short (30 cm), chair height.

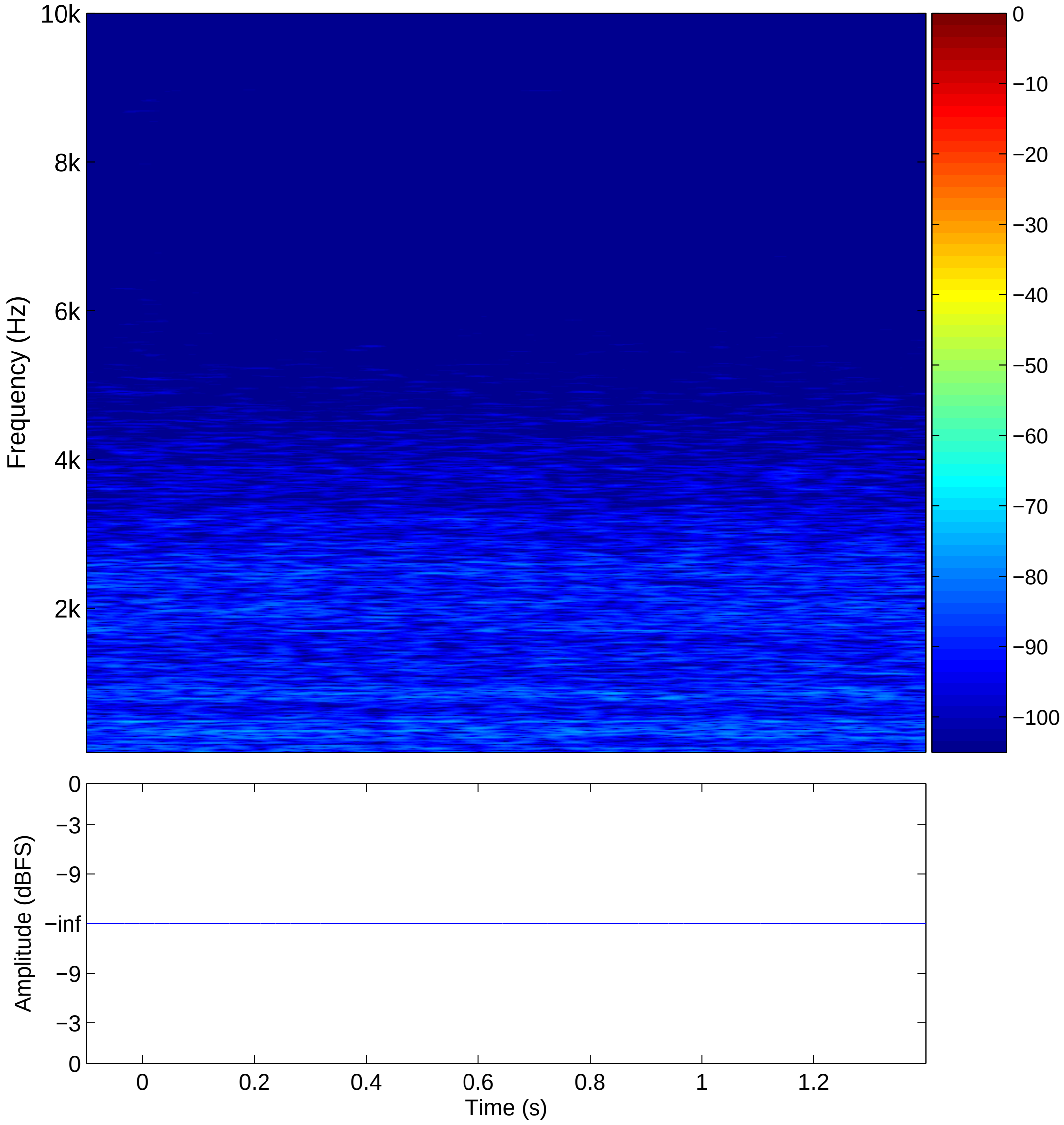


Figure C.25: Noise produced by fan that was used during test.

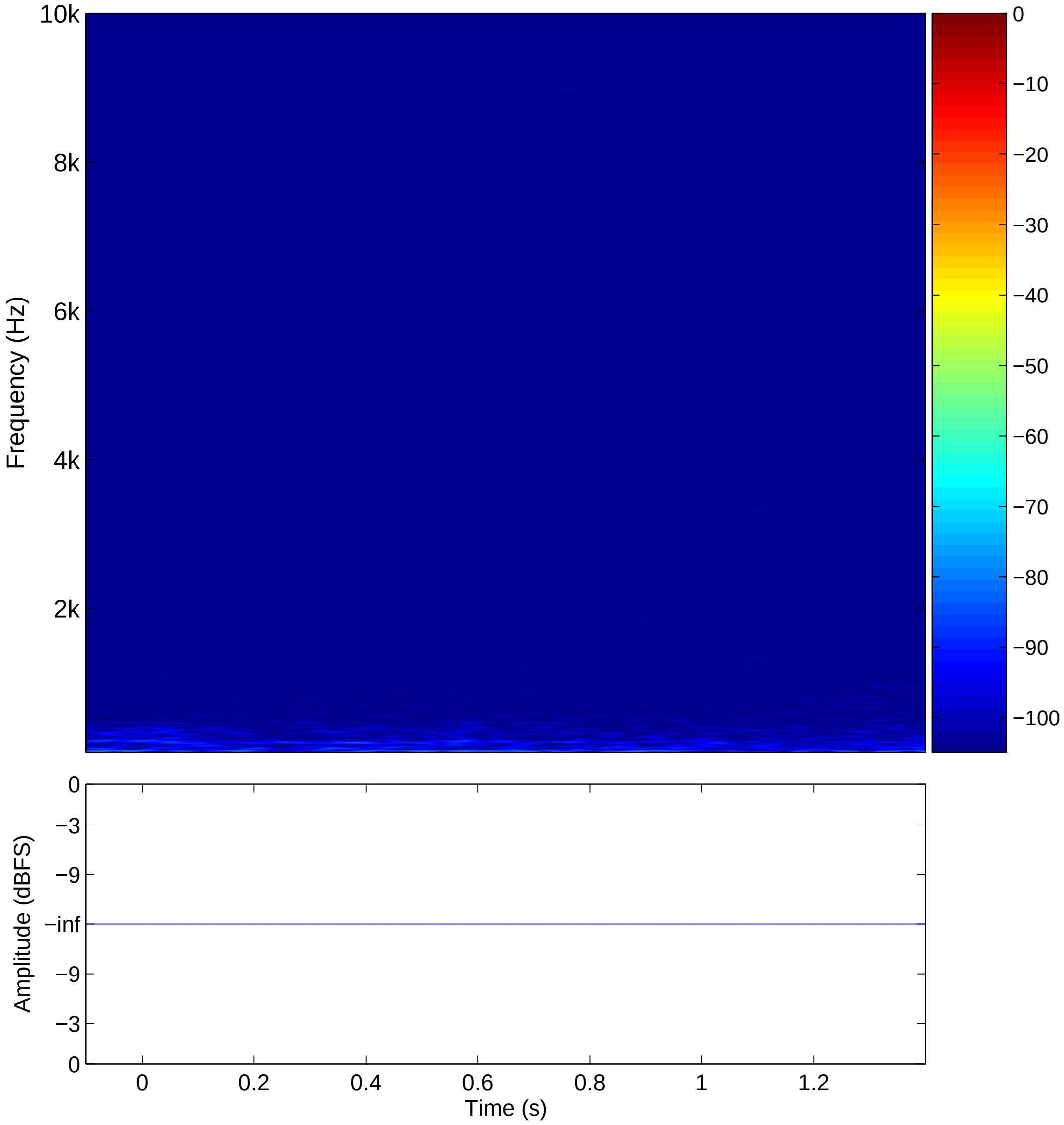


Figure C.26: Background noise without fan.

Additional Results

A few plots, not presented in Chapter 5, are presented here. They complement the results presented in Subsection 5.3.2.

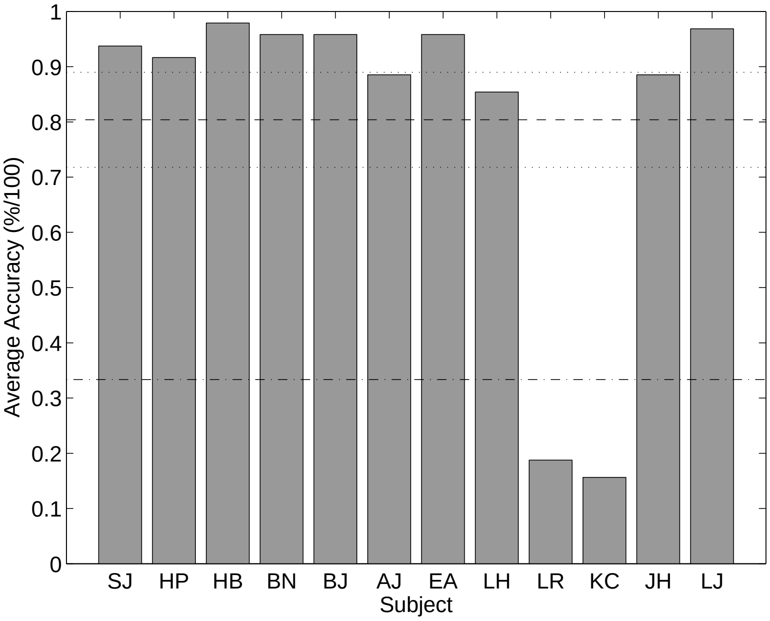


Figure C.27: Average material accuracy for unaided hearing-impaired subjects. Chance level (33.3%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

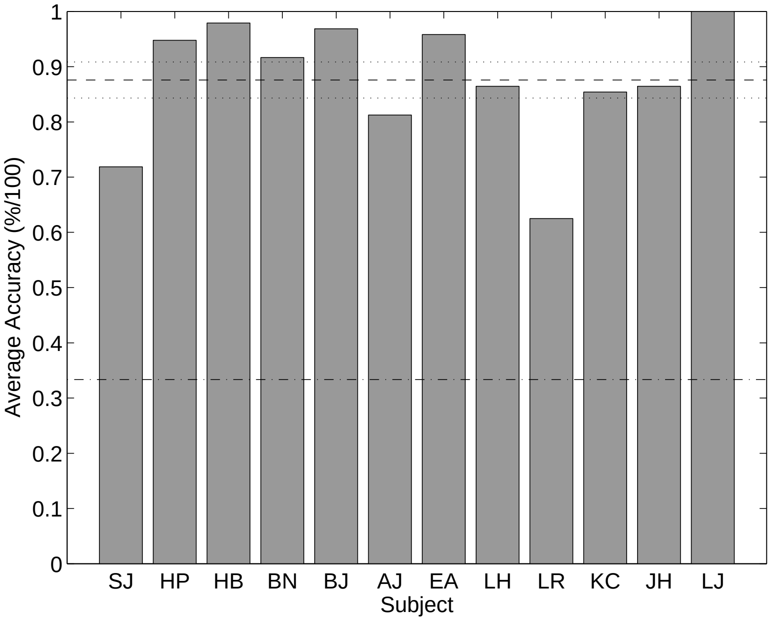


Figure C.28: Average material accuracy for aided hearing-impaired subjects. Chance level (33.3%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

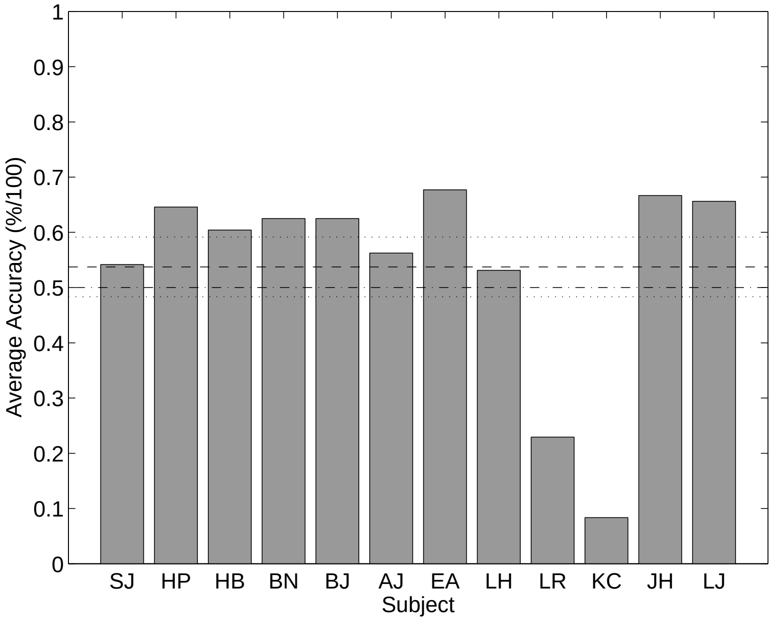


Figure C.29: Average height accuracy for unaided hearing-impaired subjects. Chance level (50%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

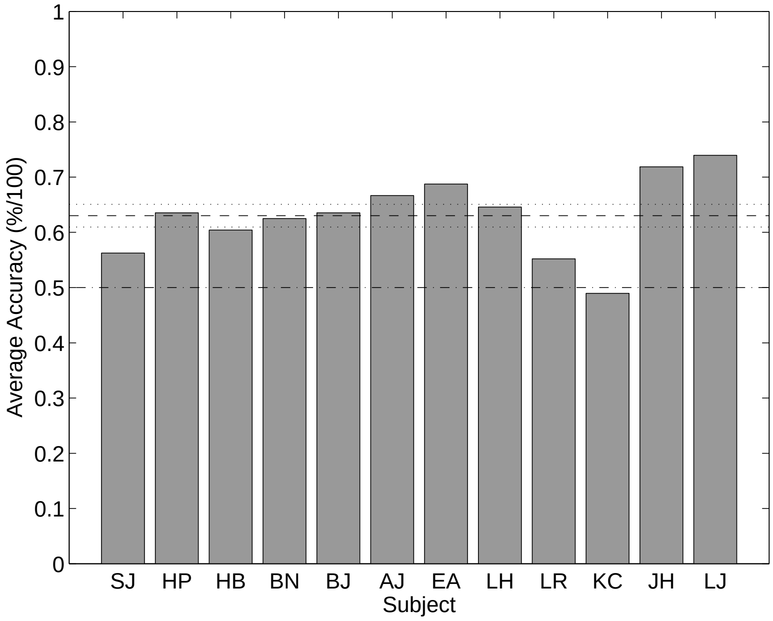


Figure C.30: Average height accuracy for aided hearing-impaired subjects. Chance level (50%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

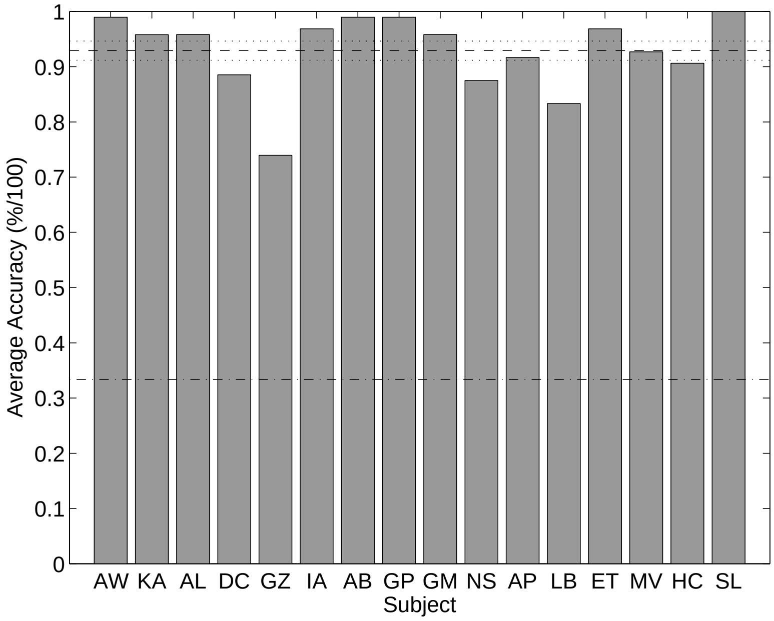


Figure C.31: Average material accuracy for unaided normal-hearing subjects. Chance level (33.3%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

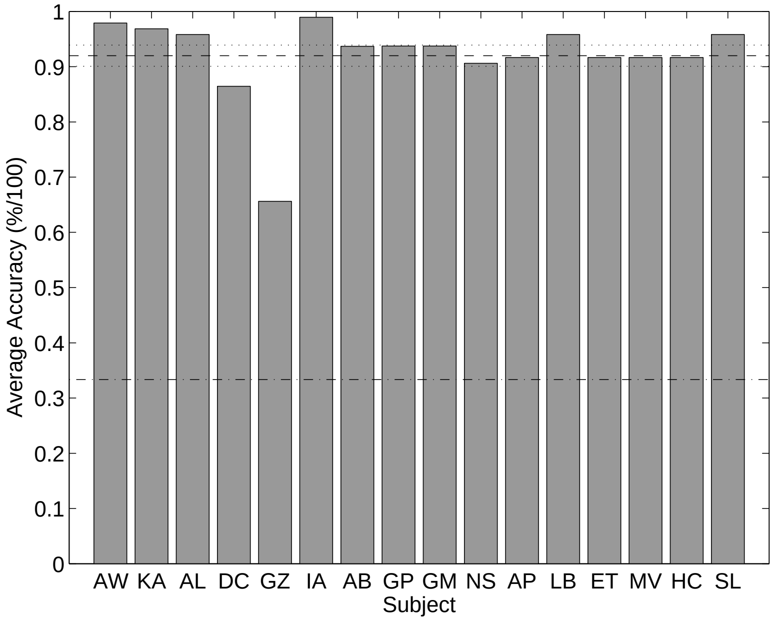


Figure C.32: Average material accuracy for aided normal-hearing subjects. Chance level (33.3%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

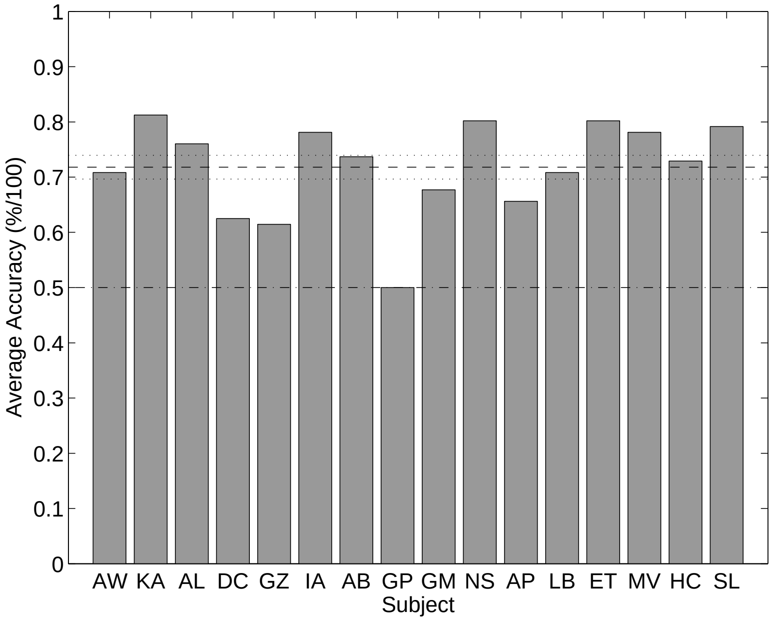


Figure C.33: Average length accuracy for aided normal-hearing subjects. Chance level (50%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

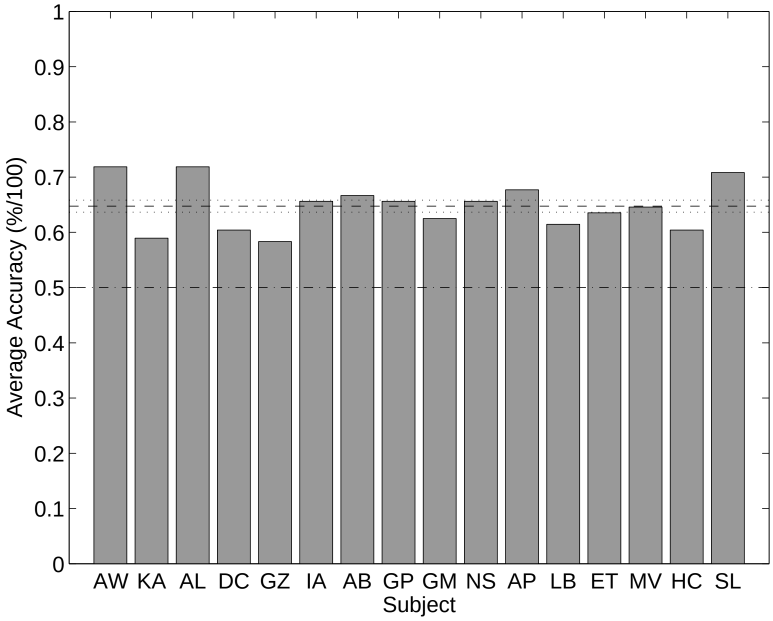


Figure C.34: Average height accuracy for unaided normal-hearing subjects. Chance level (50%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

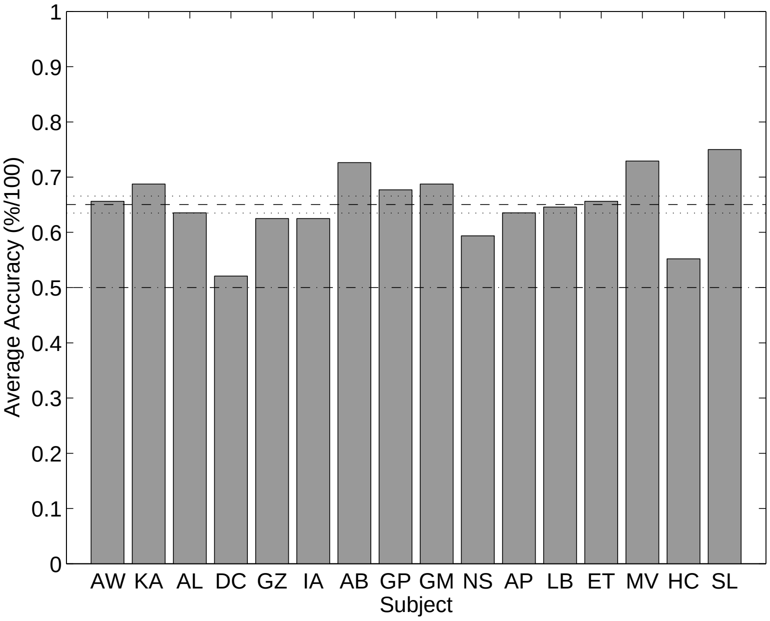


Figure C.35: Average height accuracy for aided normal-hearing subjects. Chance level (50%) is shown as the dash-dotted line. Mean of individual subject data (dashed line) is also shown with standard errors (dotted lines).

References

Anderson, K. L., Peck, A. J., & Carello, C. (1997). Rotational inertia constrains perception of object length by sound. In M. A. Schmuckler & J. M. Kennedy (Eds.), *Studies in Perception and Action IV: Ninth International Conference on Perception and Action* (pp. 79–82). Mahwah, NJ: Lawrence Erlbaum Associates.

Ashmead, D. H., Davis, D. L., & Northington, A. (1995). Contribution of listeners' approaching motion to auditory distance perception. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(2), 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*(2), 250–267.

Ballas, J. A. (2002). What is that sound? Some implications for sound design. Technical report, Naval Research Lab, Washington, D.C.

Ballas, J. A. & Mullins, T. (1991). Effects of context on the identification of everyday sounds. *Human Performance*, *4*(3), 199–219.

Batteau, D. W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *168*(1011), 158–180.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: their structure and common design principles. *Human-Computer Interaction*, *4*(1), 11–44.

Blauert, J. (1996). *Spatial Hearing: The Psychophysics of Human Sound Localization* (Revised ed.). Cambridge, Massachusetts: The MIT Press.

Bonebright, T. L. (2001). Perceptual structure of everyday sounds: A multidimensional scaling approach. In *Proceedings of the 2001 International Conference on Auditory Display*, (pp. 73–78).

Bootsma, R. J. (1989). Accuracy of perceptual processes subserving different perception-action systems. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *41*, 489–500.

Borg, E. (1998). Audiology in an ecological perspective – development of a conceptual framework. *Scandinavian Audiology*, *27*(Supplement 49), 132–139.

Bradley, J., Sato, H., & Picard, M. (2003). On the importance of early reflections for speech in rooms. *Journal of the Acoustical Society of America*, *113*(6), 3233–3244.

Brazil, E. & Fernström, M. (2003). Where's that sound? Exploring arbitrary user classifications of sounds for audio collection management. In *Proceedings of the 2003 International Conference on Auditory Display*, (pp. 66–70).

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: The MIT Press.

Buchler, M., Allegro, S., Launer, S., & Dillier, N. (2005). Sound classification in hearing aids inspired by auditory scene analysis. *ASP – EURASIP Journal on Applied Signal Processing*, *2005*(18), 2991–3002.

Cabe, P. A. & Pittenger, J. B. (2000). Human sensitivity to acoustic information from vessel filling. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 313–324.

Carello, C., Anderson, K. L., & Kunkler-Peck, A. J. (1998). Perception of object length by sound. *Psychological Science*, *9*(3), 211–214.

Carello, C., Wagman, J. B., & Turvey, M. T. (2005). Acoustic specification of object properties. In J. Anderson & B. Anderson (Eds.), *Moving Image Theory: Ecological Considerations* (pp. 79–104). Carbondale, IL: Southern Illinois University Press.

Casey, M. A. (1998). *Auditory Group Theory with Applications to Statistical Basis Methods for Structured Audio*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Çelik, E., Waye, K. P., & Møller, H. (2005). The importance of recording and playback technique for assessment of annoyance. In *Proceedings of Forum Acusticum, Budapest*, (pp. 187–192).

Cooper, M. M., Janovicz, A., & Carello, C. (2001). Length perception by sound. In G. A. Burton & R. C. Schmidt (Eds.), *Studies in Perception and Action VI* (pp. 5–8). Mahwah, NJ: Lawrence Erlbaum Associates.

Cox, R. M. & Alexander, G. C. (1999). Measuring satisfaction with amplification in daily life: The SADL scale. *Ear and Hearing*, *20*(4), 306–320.

Cox, R. M. & Alexander, G. C. (2001). Validation of the SADL questionnaire. *Ear and Hearing*, *22*(2), 151–160.

Cox, R. M., Gray, G. A., & Alexander, G. C. (2001). Evaluation of a revised speech in noise (RSIN) test. *Journal of the American Academy of Audiology*, *12*(8), 423–433.

Cusack, R. & Carlyon, R. P. (2004). Auditory perceptual organization inside and outside the laboratory. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 16–48). New York: Academic Press.

Dai, H. (1995). On measuring psychometric functions: a comparison of the constant-stimulus and adaptive up-down methods. *Journal of the Acoustical Society of America*, *98*(6), 3135–3139.

Darwin, C. (1984a). Perceiving vowels in the presence of another sound: constraints on formant perception. *Journal of the Acoustical Society of America*, *76*(6), 1636–47.

Darwin, C., Hukin, R., & Al-Khatib, B. (1995). Grouping in pitch perception: evidence for sequential constraints. *Journal of the Acoustical Society of America*, *98*(21), 880–885.

Darwin, C. J. (1984b). Auditory processing and speech perception. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of Language Processes* (pp. 197–209). Hillsdale, NJ: Lawrence Erlbaum Associates.

Darwin, C. J., Pattison, H., & Gardner, R. B. (1989). Vowel quality changes produced by surrounding tone sequences. *Perception & Psychophysics*, *45*(4), 333–342.

Dau, T., Ewert, S. D., & Oxenham, A. J. (2004). Effects of concurrent and sequential streaming in comodulation masking release. In D. Pressnitzer, A. de Cheveigné, S. McAdams, & L. Collet (Eds.), *Auditory Signal Processing: Physiology, Psychoacoustics, and Models* (pp. 335–343). New York: Springer Verlag.

Deutsch, D. (1982). Grouping mechanisms in music. In D. Deutsch (Ed.), *The Psychology of Music*. New York: Academic Press.

Dillon, H. (2001). *Hearing Aids*. Sydney: Boomerang Press.

Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *Journal of the Acoustical Society of America*, *102*(4), 2403–2411.

Doutaut, V., Matignon, D., & Chaigne, A. (1998). Numerical simulations of xylophones. II. Time-domain modeling of the resonator and of the radiated sound pressure. *Journal of the Acoustical Society of America*, *104*(31), 1633–1647.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA noises: Artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. *Audiology*, *40*(3), 148–157.

Duifhuis, H., Willems, L. F., & Sluyter, R. J. (1982). Measurement of pitch in speech: An implementation of Goldstein's theory of pitch perception. *Journal of the Acoustical Society of America*, *71*(6), 1568–1580.

Elliott, L. L. (1971). Backward and forward masking. *Audiology*, *10*, 65–76.

Fabiani, M., Kazmerski, V. A., Cycowicz, Y. M., & Friedman, D. (1996). Naming norms for brief environmental sounds: Effects of age and dementia. *Psychophysiology*, *33*(4), 462–475.

Fellgett, P. (1974). Ambisonic reproduction of directionality in surround sound systems. *Nature*, *252*, 534–538.

Femmam, S., M'Sirdi, N. K., & Ouahabi, A. (2001). Perception and characterization of materials using signal processing techniques. *IEEE Transactions on Instrumentation and Measurement*, *50*(5), 1203–1211.

Feynman, R. P. (1984). *Surely You're Joking, Mr. Feynman! Adventures of a Curious Character*. New York: W. W. Norton & Company.

Fitch, W. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*(2), 1213–1222.

Fitzpatrick, P., Carello, C., & Turvey, M. T. (1994). Eigenvalues of the inertia tensor and exteroception by the "muscular sense". *Neuroscience*, *60*(2), 551–568.

Fletcher, N. H. & Rossing, T. D. (1998). *The Physics of Musical Instruments* (Second ed.). New York: Springer-Verlag.

Fowler, C. (1990). Sound-producing sources as objects of perception: Rate normalization and nonspeech perception. *Journal of the Acoustical Society of America*, *88*(3), 1236–49.

Freed, D. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. *Journal of the Acoustical Society of America*, *87*(1), 311–22.

Gatehouse, S. (2005). Evaluating the effectiveness of hearing aid fittings in service delivery and research contexts. In Rasmussen, A. N. & Poulsen, T. (Eds.), *Twenty-First Danavox Symposium: Hearing Aid Fitting*, (pp. 423–437).

Gatehouse, S., Naylor, G., & Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *International Journal of Audiology*, *42*(Supplement 1), S77–S85.

Gaver, W. W. (1988). *Everyday Listening and Auditory Icons.* PhD thesis, University of California, San Diego, CA.

Gaver, W. W. (1993a). How do we hear in the world?: Explorations in ecological acoustics. *Ecological Psychology*, *5*(4), 285–313.

Gaver, W. W. (1993b). Synthesizing auditory icons. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems, Amsterdam, The Netherlands*, (pp. 228–235). ACM Press.

Gaver, W. W. (1993c). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Pscychology*, *5*(1), 1–29.

Gelfand, S., Ross, L., & Miller, S. (1988). Sentence reception in noise from one versus two sources: effects of aging and hearing loss. *Journal of the Acoustical Society of America*, *83*(1), 248–256.

Gelfand, S. A. (2001). *Essentials of Audiology* (Second ed.). New York: Thieme.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems.* Westport, CT: Greenwood Press.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Giordano, B. L. (2003). Material categorization and hardness scaling in real and synthetic impact sounds. In D. Rocchesso & F. Fontana (Eds.), *The Sounding Object* (pp. 73–94). Edizioni di Mondo Estremo. http://www.soundobject.org/SObBook/.

Giordano, B. L. (2005). *Sound Source Perception in Impact Sounds.* PhD thesis, Department of General Psychology, University of Padova, Italy.

Giordano, B. L. & McAdams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *Journal of the Acoustical Society of America*, *119*(2), 1171–1181.

Giordano, B. L. & Petrini, K. (2003). Hardness recognition in synthetic sounds. In *Proceedings of the Stockholm Music Acoustics Conference*, (pp. 611–614).

Goldhor, R. (1993). Recognition of environmental sounds. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 1993*, *1*, 149–152.

Gordon, M. S. & Rosenblum, L. D. (2001). Perception of acoustic occlusion using body-scaled judgements. In G. A. Burton & R. C. Schmidt (Eds.), *Studies in Perception and Action VI* (pp. 9–12). Mahwah, NJ: Lawrence Erlbaum Associates.

Gordon, M. S. & Rosenblum, L. D. (2004). Perception of sound-obstructing surfaces using body-scaled judgments. *Ecological Psychology*, *16*(2), 87–114.

Grassi, M. (2002). Recognising the size of objects from sounds with manipulated acoustical parameters. In Da Silva, J. A., Matsushima, E. H., & Ribeiro-Filho, N. P. (Eds.), *Proceedings of the 18th Annual Meeting of the International Society for Psychophysics, Rio de Janeiro, Brasil*, (pp. 392–397). The International Society for Psychophysics, Fechner Day.

Grassi, M. (2003). *Two Researches in Ecological Psychoacoustics*. PhD thesis, University of Padova, Italy.

Grassi, M. (2005). Do we hear size or sound? Balls dropped on plates. *Perception & Psychophysics*, *67*(2), 274–284.

Grassi, M. & Burro, R. (2003). Impact sounds. In D. Rocchesso & F. Fontana (Eds.), *The Sounding Object* (pp. 47–72). Edizioni di Mondo Estremo. http://www.soundobject.org/SObBook/.

Green, P. (2004). Missing links between auditory scene analysis and automatic speech recognition. In *Proceedings of the Eighth International Congress on Acoustics*, (pp. 1153–1159).

Griffin, D. R. (1959). *Echos of Bats and Men*. New York: Anchor Books.

Guastavino, C., Katz, B. F. G., Polack, J.-D., Levitin, D. J., & Dubois, D. (2005). Ecological validity of soundscape reproduction. *Acta Acustica united with Acustica*, *91*(2), 333–341.

Guski, R. (1992). Acoustic tau: An easy analogue to visual tau? *Ecological Psychology*, *4*(3), 189–197.

Gygi, B. (2001). *Factors in the Identification of Environmental Sounds.* PhD thesis, Indiana University, Bloomington, IN.

Gygi, B., Kidd, G. R., & Watson, C. S. (1999). Environmental sounds: Detection, identification and similarity. In *NIPS*99, Workshop on Neural Mechanisms of Music Processing.*

Gygi, B., Kidd, G. R., & Watson, C. S. (2000). Identification and similarity judgments of environmental sounds. *Journal of the Acoustical Society of America*, *107*(2), 2820.

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*, *115*(3), 1252–1265.

Gygi, B., Watson, C., & Kidd, G. (2000). Effects of low-pass and high-pass filtering on the identification of environmental sounds. *Journal of the Acoustical Society of America*, *108*(52), 2642.

Halpern, D. L., Blake, R., & Hillenbrand, J. (1986). Psychoacoustics of a chilling sound. *Perception & Psychophysics*, *39*(2), 77–80.

Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events.* Cambridge, MA: The MIT Press.

Haykin, S. & Chen, Z. (2005). The cocktail party problem. *Neural Computation*, *17*(9), 1875–1902.

Heft, H. (1993). A methodological note on overestimates of reaching distance: Distinguishing between perceptual and analytical judgments. *Ecological Psychology*, *5*(3), 255–271.

Heine, W.-D. & Guski, R. (1991). Listening: The perception of auditory events? *Ecological Psychology*, *3*, 263–275.

Heine, W.-D., Guski, R., & Pittenger, J. B. (1993a). Auditory perception of numbers. In S. S. Valenti & J. B. Pittenger (Eds.), *Studies in Perception and Action II: Posters Presented at the VIIth International Conference on Event Perception and Action : August 8-13, 1993, University of British Columbia, Vancouver* (pp. 259–262). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heine, W.-D., Guski, R., & Pittenger, J. B. (1993b). Can listeners catch approaching sound sources? In S. S. Valenti & J. B. Pittenger (Eds.), *Studies in Perception and Action II: Posters Presented at the VIIth International Conference on Event Perception and Action : August 8-13, 1993, University of*

British Columbia, Vancouver (pp. 238–241). Hillsdale, NJ: Lawrence Erlbaum Associates.

Heine, W.-D., Guski, R., & Pittenger, J. B. (1993c). Perceiving numbers of steel balls by audition. In Schick, A. (Ed.), *Contributions to Psychological Acoustics: Results of the Sixth Oldenburg Symposium on Psychological Acoustics*, (pp. 361–371). Bibliotheks- und Informationssystem der Universität Oldenburg.

Heller, L. M. (2002a). When sound effects are better than the real thing. http://www.acoustics.org/press/143rd/Heller.html. Popular version of paper 1pPP10.

Heller, L. M. (2002b). When sound effects are better than the real thing. *Journal of the Acoustical Society of America*, *111*(5), 2339. 1pPP10 - Abstract.

Houben, M., Kohlrausch, A., & Hermes, D. (2004). Perception of the size and speed of rolling balls by sound. *Speech Communication*, *43*(4), 331–345.

Houben, M. M., Kohlrausch, A., & Hermes, D. J. (2001). Auditory cues determining the perception of size and speed of rolling balls. In *Proceedings of the 2001 International Conference on Auditory Display, Espoo, Finland, July 29-August 1, 2001*, (pp. 105–110).

Houben, M. M. J. (2002). *The Sound of Rolling Objects: Perception of Size and Speed*. PhD thesis, Technische Universiteit Eindhoven.

Houix, O., McAdams, S., & Caussé, R. (1999). Auditory categorization of sound sources. In M. A. Grealy & J. A. Thomson (Eds.), *Studies in Perception and Action V: Tenth International Conference on Perception and Action* (pp. 47–51). Mahwah, NJ: Lawrence Erlbaum Associates.

Ishibashi, M. & Preis, A. (2005a). Auditory cue for perception of weight of falling balls. *Archives of Acoustics*, *30*(4, Supplement), 137–142.

Ishibashi, M. & Preis, A. (2005b). Psychoacoustic estimation of differential mass threshold. In *Proceedings of Forum Acusticum, Budapest*, (pp. 1671–1675).

ISO 8253-1 (1989). ISO 8253-1 International Standard: Acoustics – Audiometric test methods – Part 1: Basic pure tone air and bone conduction threshold audiometry. Technical report, International Organization for Standardization.

Ito, K., Stoffregen, T. A., Donohue, S., & Nelson, G. K. (2001). Blind adults exhibit postural responses to a moving room. In G. A. Burton & R. C. Schmidt (Eds.), *Studies in Perception and Action VI* (pp. 1–4). Mahwah, NJ: Lawrence Erlbaum Associates.

Janku, L. (2004). Artificial perception: Auditory decomposition of mixtures of environmental sounds – combining information theoretical and supervised pattern recognition approaches. In P. van Emde Boas, J. Pokorný, M. Bieliková, & J. Stuller (Eds.), *Proceedings of SOFSEM 2004: Theory and Practice of Computer Science. Thirtieth Conference on Current Trends in Theory and Practice of Computer Science (Lecture Notes in Computer Science Vol. 2932)* (pp. 229–40). ERCIM, Microsoft, Deloitte & Touche, SOFTEC Bratislava, Centrum.cz.

Jenison, R. L. (1997). On acoustic information for motion. *Ecological Psychology*, *9*(2), 131–151.

Jenkins, J. J. (1985). Acoustic information for objects, places, and events. In W. H. Warren, Jr. & R. E. Shaw (Eds.), *Persistence and Change: Proceedings of the First International Conference on Event Perception* (pp. 115–138). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jesteadt, W., Wier, C. C., & Green, D. M. (1977). Intensity discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, *61*, 169–177.

Johansson, G. (1985). About visual event perception. In W. H. Warren, Jr. & R. E. Shaw (Eds.), *Persistence and Change: Proceedings of the First International Conference on Event Perception* (pp. 29–54). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kettles, N. (2006). Can you hear yourself above the din? Or has your voice been drowned out by the cacophony of society straining at the seams? *The Ecologist*, *36*(3), 42–46.

Kiang, N. Y.-S., Watanabe, T., Thomas, E. C., & Clark, L. F. (1965). *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve.* Cambridge, Massachusetts: MIT.

Kim, C., Stoffregen, T. A., Ito, K., & Bardy, B. G. (2005). Coupling of movement to acoustic flow in sighted adults. In H. Heft & K. L. Marsh (Eds.), *Studies in Perception and Action VIII* (pp. 169–172). Mahwah, NJ: Lawrence Erlbaum Associates.

Kirkwood, B. C. (2005a). As good as the real thing? Performance differences for live versus recorded stimuli in an everyday listening task. In Rasmussen, A. N. & Poulsen, T. (Eds.), *Twenty-First Danavox Symposium: Hearing Aid Fitting*, (pp. 519–529).

Kirkwood, B. C. (2005b). As good as the real thing? Performance differences for live versus recorded stimuli in an everyday listening task. Poster Presented at the Twenty-First Danavox Symposium: Hearing Aid Fitting, Kolding, Denmark, August 31–September 2, 2005.

Kirkwood, B. C. (2005c). The influence of presentation method on auditory length perception. In H. Heft & K. L. Marsh (Eds.), *Studies in Perception and Action VIII* (pp. 145–148). Mahwah, NJ: Lawrence Erlbaum Associates.

Kirkwood, B. C. (2005d). The influence of presentation method on auditory length perception. Poster Presented at the Thirteenth International Conference on Perception and Action, Monterey, CA, July 5-10, 2005.

Kirkwood, B. C. (2005e). Maintaining realism in auditory length-perception experiments. In Stoustrup, J. (Ed.), *PhD Papers in Technology and Science - Fall 2005: Selected Papers from the PhD course "Writing and Reviewing Scientific Papers"*, (pp. 46–48). Aalborg University, Denmark.

Klatzky, R., Pai, D., & Krotkov, E. (2000). Perception of material from contact sounds. *Presence*, *9*(4), 399–410.

Krotkov, E., Klatzky, R., & Zumel, N. (1996). Robotic perception of material: Experiments with shape-invariant acoustic measures of material type. In J. S. Oussama Khatib (Ed.), *Experimental Robotics IV: Lecture notes in control and information science* (pp. 204–211). New York: Springer Verlag.

Kunkler-Peck, A. J. & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 279–294.

Kürer, R., Plenge, G., & Wilkens, H. (1969). Correct spatial sound perception rendered by a special 2-channel recording method. Presented at the 37th Convention of the Audio Engineering Society October 13-16, 1969. Audio Engineering Society Preprint Number 666.

Lakatos, S., McAdams, S., & Caussé, R. (1997). The representation of auditory source characteristics: Simple geometric form. *Perception & Psychophysics*, *59*(8), 1180–1190.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5*, 437–459.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*(2), 467–77.

Li, X., Logan, R., & Pastore, R. (1991). Perception of acoustic source characteristics: Walking sounds. *Journal of the Acoustical Society of America*, *90*(6), 3036–49.

Loomis, J. M., Fujita, N., Da Silva, J. A., & Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 906–921.

Lunner, T. (2003). Cognitive function in relation to hearing aid use. *International Journal of Audiology*, *42*(Supplement 1), S49–S58.

Lutfi, R. (2001). Auditory detection of hollowness. *Journal of the Acoustical Society of America*, *110*(2), 1010–19.

Lutfi, R. & Oh, E. (1997). Auditory discrimination of material changes in a struck-clamped bar. *Journal of the Acoustical Society of America*, *102*(6), 3647–56.

Lutfi, R., Oh, E., Storm, E., & Alexander, J. (2005). Classification and identification of recorded and synthesized impact sounds by practiced listeners, musicians, and nonmusicians. *Journal of the Acoustical Society of America*, *118*(1), 393–404.

Lutfi, R. A. & Oh, E. (1994). Auditory discrimination based on the physical dynamics of a tuning fork. *Journal of the Acoustical Society of America*, *95*(5), 2967.

Mace, W. (2005). How many invariants are there? (More than you thought). In Heft, H., Marsh, K. L., & Rosenblum, L. D. (Eds.), *Proceedings of the Thirteenth International Conference on Perception and Action, July 5-10, 2005, Monterey, CA.* (p. 14).

Marcell, M. M., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology*, *22*(6), 830–864.

McAdams, S. & Bigand, E. (Eds.). (1993). *Thinking In Sound: The Cognitive Psychology of Human Audition*. Oxford: Oxford University Press.

McAdams, S., Chaignec, A., & Roussarie, V. (2004). The psychomechanics of simulated sound sources: Material properties of impacted bars. *Journal of the Acoustical Society of America*, *115*(3), 1306–1320.

McGurk, H. & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1-2), 36–47.

Miller, J. D. & Tanis, D. C. (1971). Recognition memory for common sounds. *Psychonomic Science*, *23*(4), 307–308.

Minnaar, P., Olesen, S. K., Christensen, F., & Møller, H. (2001). Localization with binaural recordings from artificial and human heads. *AES: Journal of the Audio Engineering Society*, *49*(5), 323–336.

Møller, H. (1992). Fundamentals of binaural technology. *Applied Acoustics*, *36*, 171–218.

Møller, H., Hammershøi, D., Jensen, C. B., & Sørensen, M. F. (1999). Evaluation of artificial heads in listening tests. *AES: Journal of the Audio Engineering Society*, *47*(3), 83–100.

Moore, B., Glasberg, B., & Peters, R. (1985). Relative dominance of individual partials in determining the pitch of complex tones. *Journal of the Acoustical Society of America*, *77*(5), 1853–60.

Moore, B. C. J. (2003). *An Introduction to the Psychology of Hearing* (Fifth ed.). San Diego, CA: Academic Press.

Narins, P. M. & Smith, S. L. (1986). Clinal variation in anuran advertisement calls: Basis for acoustic isolation? *Behavioral Ecology and Sociobiology*, *19*, 135–141.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123–124.

Neuhoff, J. G. (2001a). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, *13*(2), 87–110.

Neuhoff, J. G. (2001b). The audible facing angle. *Acoustics Research Letters Online*, *2*(4), 109–114.

Neuhoff, J. G. (2001c). Perceiving acoustic source orientation with dynamic rotation cues. In G. A. Burton & R. C. Schmidt (Eds.), *Studies in Perception and Action VI* (pp. 13–16). Mahwah, NJ: Lawrence Erlbaum Associates.

Neuhoff, J. G. (2003). Twist and shout: Audible facing angles and dynamic rotation. *Ecological Psychology*, *15*(4), 335–352.

Neuhoff, J. G. (2004a). Auditory motion and localization. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 87–111). New York: Academic Press.

Neuhoff, J. G. (Ed.). (2004b). *Ecological Psychoacoustics.* Academic Press.

Neuhoff, J. G. (2004c). Ecological psychoacoustics: Introduction and history. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 1–13). New York: Academic Press.

Neuhoff, J. G. (2004d). Interacting perceptual dimensions. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 249–269). New York: Academic Press.

Noble, W. (1983). Hearing, hearing impairment, and the audible world: A theoretical essay. *Audiology*, *22*, 325–338.

Noble, W. & Hétu, R. (1994). An ecological approach to disability and handicap in relation to impaired hearing. *Audiology*, *33*, 117–126.

Okura, M. (1999). How to pick up the defective canned foods by tapping. In *Proceedings of the Tenth International Conference on Perception and Action, August 8-14, 1999, University of Edinburgh, Scotland.* (p. 85).

Oxenham, A. J. & Moore, B. C. J. (1994). Modeling the additivity of nonsimultaneous masking. *Hearing Research*, *80*, 105–118.

Palmer, C., Nelson, C., & Lindley, G. (1998). The functionally and physiologically plastic adult auditory system. *Journal of the Acoustical Society of America*, *103*(4), 1705–1721.

Papakonstantinou, A. (2005). Relation between speech intelligibility, temporal auditory perception, and auditory-evoked potentials. Master's thesis, Technical University of Denmark, Kongens Lyngby, Denmark.

Pearsall, J. (Ed.). (1999). *The Concise Oxford Dictionary* (Tenth ed.). Oxford, United Kingdom: Oxford University Press.

Peltonen, V. T. K., Eronen, A. J., Parviainen, M. P., & Klapuri, A. P. (2001). Recognition of everyday auditory scenes: Potentials, latencies and cues. In *Proceedings of the 110th Audio Engineering Society Convention.*

Pichora-Fuller, M. K. (1999). Acoustic ecology: Concept and case study. *Canadian Acoustics*, *27*(3), 120.

Pickles, J. O. (1988). *An Introduction to the Physiology of Hearing* (Second ed.). London: Academic Press.

Pittenger, J., Jordan, J., Belden, A., Goodspeed, P., & Brown, F. (1997). Auditory and haptic information support perception of size. In M. A. Schmuckler & J. M. Kennedy (Eds.), *Studies in Perception and Action IV: Ninth International Conference on Perception and Action* (pp. 103–105). Mahwah, NJ: Lawrence Erlbaum Associates.

Pittenger, J. B. & Mincy, M. D. (1999). Haptic and auditory information support perception of size: Fine granules. In M. A. Grealy & J. A. Thomson (Eds.), *Stud-*

ies in Perception and Action V: Tenth International Conference on Perception and Action (pp. 68–72). Mahwah, NJ: Lawrence Erlbaum Associates.

Plomp, R. (1977). Acoustical aspects of cocktail parties. *Acustica*, *38*(3), 186–191.

Plomp, R. (2002). *The Intelligent Ear: On The Nature of Sound Perception*. Mahwah, NJ: Lawrence Erlbaum Associates.

Popper, A. N. & Fay, R. R. (1997). Evolution of the ear and hearing: Issues and questions. *Brain, Behavior and Evolution*, *50*, 213–221.

Poulsen, T. & Keidser, G. (1991). Speech recognition in young and elderly normal-hearing listeners in a closed response test. *Scandinavian Audiology*, *20*, 245–249.

Pralong, D. & Carlile, S. (1996). The role of individualized headphone calibration for the generation of high fidelity virtual auditory space. *Journal of the Acoustical Society of America*, *100*(6), 3785–3793.

Proffitt, D., Stefanucci, J., Banton, T., & Epstein, W. (2003). The role of effort in distance perception. *Psychological Science*, *14*, 106–113.

Proops, D. W., Donaldson, I., Cooper, H. R., Thomas, J., Burrell, S. P., Stoddart, R. L., Moore, A., & Cheshire, I. M. (1999). Outcomes from adult implantation, the first 100 patients. *Journal of Laryngology and Otology*, *113*, 5–13.

Ramsdell, D. A. (1978). The psychology of the hard-of-hearing and the deafened adult. In H. Davis & S. R. Silverman (Eds.), *Hearing and Deafness* (pp. 499–510). New York: Holt, Rinehart & Winston.

Reed, C. M. & Delhorne, L. A. (2005). Reception of environmental sounds through cochlear implants. *Ear and Hearing*, *26*(1), 48–61.

Repp, B. (1987). The sound of two hands clapping: an exploratory study. *Journal of the Acoustical Society of America*, *81*(4), 1100–9.

Rieser, J. J., Ashmead, D. H., Talor, C. R., & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, *19*, 675–689.

Robart, R. L. & Rosenblum, L. D. (2005a). Hearing silent shapes: Identifying the shape of a sound-obstructing surface. In H. Heft & K. L. Marsh (Eds.), *Studies in Perception and Action VIII* (pp. 149–152). Mahwah, NJ: Lawrence Erlbaum Associates.

Robart, R. L. & Rosenblum, L. D. (2005b). Hearing space: Identifying rooms by reflected sound. In H. Heft & K. L. Marsh (Eds.), *Studies in Perception and Action VIII* (pp. 153–156). Mahwah, NJ: Lawrence Erlbaum Associates.

Roberts, B. & Moore, B. (1991). The influence of extraneous sounds on the perceptual estimation of first-formant frequency in vowels under conditions of asynchrony. *Journal of the Acoustical Society of America*, *89*(6), 2922–32.

Rolland, J., Meyer, C., Arthur, K., & Rinalducci, E. (2002). Method of adjustments versus method of constant stimuli in the quantification of accuracy and precision of rendered depth in head-mounted displays. *Presence*, *11*(6), 610–25.

Rosenblum, L. D. (2004). Perceiving articulatory events: Lessons for an ecological psychoacoustics. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 219–248). New York: Academic Press.

Rosenblum, L. D. & Fowler, C. A. (1991). Audiovisual investigation of the loudness-effort effect for speech and nonspeech events. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 976–985.

Rosenblum, L. D., Gordon, M. S., & Jarquin, L. (2000). Echolocating distance by moving and stationary listeners. *Ecological Psychology*, *12*(3), 181–206.

Rosenblum, L. D., Wuestefeld, A. P., & Anderson, K. L. (1996). Auditory reachability: An affordance approach to the perception of sound source distance. *Ecological Psychology*, *8*(1), 1–24.

Rosenblum, L. D., Wuestefeld, A. P., & Saldaña, H. M. (1993). Auditory looming perception: Influences on anticipatory judgements. *Perception*, *22*, 1467–1482.

Rossing, T. D. (1990). *The Science of Sound* (Second ed.). New York: Addison-Wesley.

Roussarie, M. V. (1999). *Analyse perceptive de structures vibrantes simulees par modele physique*. PhD thesis, Université du Maine, France.

Russell, M. K. (1997). Acoustic perception of sound source occlusion. In M. A. Schmuckler & J. M. Kennedy (Eds.), *Studies in Perception and Action IV: Ninth International Conference on Perception and Action* (pp. 92–94). Mahwah, NJ: Lawrence Erlbaum Associates.

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*(3), 223–237.

Russell, M. K. & Schuler, K. (2001). Auditory perception of distance in 3-dimensional space. In G. A. Burton & R. C. Schmidt (Eds.), *Studies in Perception and Action VI* (pp. 17–20). Mahwah, NJ: Lawrence Erlbaum Associates.

Saldaña, H. M. & Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Perception & Psychophysics*, *54*(3), 406–416.

Sandvad, J. (1999). Auditory perception of reverberant surroundings. *Journal of the Acoustical Society of America*, *105*(2), 1193.

Schaeffer, P. (1966). *Traité des Objets Musicaux.* Paris: Seuil.

Schmuckler, M. A. (2004). Pitch and pitch structures. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 271–315). New York: Academic Press.

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology*, *56*(3), 139–152.

Shafiro, V. (2004a). Acoustic correlates of the variation in the number of channels needed to identify the sources of familiar environmental sounds. Presentation at the 2004 Midwinter Research Meeting of the Association for Research in Otolaryngology.

Shafiro, V. (2004b). *Perceiving the Sources of Environmental Sounds with a Varying Number of Spectral Channels.* PhD thesis, City University of New York.

Shafiro, V. & Gilichinskaya, Y. (2004). The role of spectral asynchrony in the identification of spectrally smeared environmental sounds. *Journal of the Acoustical Society of America*, *115*(5), 2459.

Shafiro, V. & Gygi, B. (2004). How to select stimuli for environmental sound research and where to find them. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 590–8.

Shaw, B. K., McGowan, R. S., & Turvey, M. T. (1991). An acoustic variable specifying time-to-contact. *Ecological Psychology*, *3*(3), 253–261.

Siegel, S. & Castellan, Jr., N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences* (Second ed.). Statistics Series. New York: McGraw-Hill International Editions.

Smith, D. R. R., Patterson, R. D., Turner, R., Kawahara, H., & Irino, T. (2005). The processing and perception of size information in speech sounds. *Journal of the Acoustical Society of America*, *117*(1), 305–318.

Soli, S. D. & Nilsson, M. (1994). Assessment of communication handicap with the HINT. *Hearing Instruments*, *45*, 12–16.

Solomon, H. Y. & Turvey, M. T. (1988). Haptically perceiving the distances reachable with hand-held objects. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 404–427.

Spanik, C. & Pichora-Fuller, K. (1999). Identification of gated environmental sounds. *Canadian Acoustics*, *27*(3), 122–123.

Supa, M., Cotzin, M., & Dallenbach, K. M. (1944). Facial vision: The perception of obstacles by the blind. *American Journal of Psychology*, *57*, 133–183.

Takada, M., Tanaka, K., & Iwamiya, S. (2006). Relationships between auditory impressions and onomatopoeic features for environmental sounds. *Acoustical Science and Technology*, *27*(2), 67–79.

Tucker, S. (2003). *An Ecological Approach to the Classification of Transient Underwater Acoustic Events: Perceptual Experiments and Auditory Models*. PhD thesis, University of Sheffield, United Kingdom.

Tucker, S. & Brown, G. J. (2002a). Investigating the perception of the size, shape and material of damped and free vibrating plates. Technical report, University of Sheffield, United Kingdom. Technical Report CS-02-10.

Tucker, S. & Brown, G. J. (2002b). Perceiving the size, shape and material of struck plates. Poster presented at the British Society of Audiology Short Papers Meeting on Experimental Studies of Hearing And Deafness, Sheffield, United Kingdom, 16-17th September.

Tucker, S. & Brown, G. J. (2002c). Perceiving the size, shape and material of struck plates. *International Journal of Audiology*, *42*(6), 356.

Tucker, S. & Brown, G. J. (2003). Modelling the auditory perception of size, shape and material: Applications to the classification of transient sonar sounds. In *114th Audio Engineering Society Convention, March 22-25, Amsterdam, The Netherlands*.

Turner, C. W., Humes, L. E., Bentler, R. A., & Cox, R. M. (1996). A review of past research on changes in hearing aid benefit over time. *Ear and Hearing*, *17*(3, Supplement), S14–S25.

Turvey, M. T. (1986). The ecological approach to perceiving-acting: A pictorial essay. *Acta Psychologica*, *63*, 133–155.

Tye-Murray, N., Tyler, R. S., Woodworth, G. G., & Gantz, B. J. (1992). Performance over time with a nucleus or ineraid cochlear implant. *Ear and Hearing*, *13*, 200–209.

van den Doel, K., Kry, P. G., & Pai, D. K. (2001). FoleyAutomatic: Physically-based sound effects for interactive simulation and animation. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, (pp. 537–544).

van den Doel, K. & Pai, D. K. (1998). The sounds of physical shapes. *Presence: Teleoperators and Virtual Environments*, *7*(4), 382–395.

van den Doel, K., Pai, D. K., Adam, T., Kortchmar, L., & Pichora-Fuller, K. (2002). Measurements of perceptual quality of contact sound models. In *Proceedings of the International Conference on Auditory Display, 2002, Kyoto, Japan*, (pp. 345–349).

Van Heuven, V. J. & Van Den Broecke, M. P. R. (1979). Auditory discrimination of rise and decay times in tone and noise bursts. *Journal of the Acoustical Society of America*, *66*, 1308–1315.

VanDerveer, N. J. (1979). *Ecological Acoustics: Human Perception of Environmental Sounds*. PhD thesis, Cornell University, Ithaca, NY. Dissertation Abstracts International 40: 4543B (University Microfilms No. 8004002).

Vicente, K. J. (2003). Beyond the lens model and direct perception: Toward a broader ecological psychology. *Ecological Psychology*, *15*(3), 241.

von Helmholtz, H. L. F. (1954). *On the Sensations of Tones as the Physiological Basis for the Theory of Music*. New York: Dover. Translation by A. J. Ellis. Original work published in 1863.

Wagman, J. B. (2003a). What is the sound of one rod dropping? Poster Presented at the Twelfth International Conference on Perception and Action, Gold Coast, Australia, July 13-18, 2003.

Wagman, J. B. (2003b). What is the sound of one rod dropping? In S. Rogers & J. Effken (Eds.), *Studies in Perception and Action VII* (pp. 195–198). Mahwah, NJ: Lawrence Erlbaum Associates.

Wagman, J. B., Hopkins, K. M., & Minarik, J. L. (2005a). Does length sound like what length feels like? In H. Heft & K. L. Marsh (Eds.), *Studies in Perception and Action VIII* (pp. 162–165). Mahwah, NJ: Lawrence Erlbaum Associates.

Wagman, J. B., Hopkins, K. M., & Minarik, J. L. (2005b). Does length sound like what length feels like? Poster Presented at the Thirteenth International Conference on Perception and Action, Monterey, CA, July 5-10, 2005.

Walker, B. N. & Kramer, G. (2004). Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 149–174). New York: Academic Press.

Warren, W. H. & Whang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 371–383.

Warren, Jr., W. H., Kim, E. E., & Husney, R. (1987). The way the ball bounces: visual and auditory perception of elasticity and control of the bounce pass. *Perception*, *16*, 309–336.

Warren, Jr., W. H. & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 704–712.

Weisser, A. (2004). Sound quality in small rooms. Master's thesis, Technical University of Denmark, Kongens Lyngby, Denmark.

Werner, L. A. & Liebold, L. J. (2004). Ecological developmental psychoacoustics. In J. G. Neuhoff (Ed.), *Ecological Psychoacoustics* (pp. 191–217). New York: Academic Press.

Wier, C. C., Jesteadt, W., & Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, *61*, 178–184.

Wildes, R. P. & Richards, W. A. (1988). Recovering material properties from sound. In W. Richards (Ed.), *Natural Computation* (pp. 356–363). The MIT Press.

Witt, J. K., Proffitt, D. R., & Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception*, *33*(5), 577–590.

Witt, J. K., Proffitt, D. R., & Epstein, W. (2005). Tool use affects perceived distance, but only when you intend to use it. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 880–887.

Wrightson, K. (2000). An introduction to acoustic ecology. *Soundscape: The Journal of Acoustic Ecology*, *1*(1), 10–13.

Wuestefeld, A. P. & Rosenblum, L. D. (1993). Hearing what is reachable. In S. S. Valenti & J. B. Pittenger (Eds.), *Studies in Perception and Action II: Posters presented at the VIIth International Conference on Event Perception and Action* (pp. 254–258). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zeng, F.-G., Kong, Y.-Y., Michalewski, H. J., & Starr, A. (2005). Perceptual consequences of disrupted auditory nerve activity. *Journal of Neurophysiology*, *93*(6), 3050–3063.

Zwicker, E. & Fastl, H. (1999). *Psychoacoustics: Facts and Models* (Second ed.). Berlin: Springer Verlag.