COMMUNICATION WITHOUT WORDS: UNDERSTANDING THE IMPLICATIONS OF TEMPORAL STRUCTURE FOR AUDITORY PERCEPTION

By JESSICA GILLARD, B.A. (hons.)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2014)
McMaster University
Psychology, Neuroscience & Behaviour

TITLE: Communication without words: Understanding the implications of temporal structure for auditory perception

AUTHOR: Jessica Gillard, B.A. (hons.) (McMaster University)

SUPERVISOR: Dr. Michael Schutz

NUMBER OF PAGES: ix, 86

Preface

The following thesis consists of two manuscripts intended for publication in scientific journals. Chapter 2 contains a manuscript entitled 'Composing alarms: Considering the musical aspects of alarm design' that is currently ready for submission. The author of the present thesis is the primary author of this work, and was responsible for experimental design, data collection, entry, and analysis, and manuscript preparation. The thesis supervisor is the second author of this paper.

Chapter 3 consists of an article entitled 'Classifying the properties of sounds used in auditory perception research', which will be submitted for publication in the near future. The author of the present thesis is the primary author of this work, and was responsible for meta-data collection, entry, and analysis, the development of several R functions used for data analysis and visualization (see Chapter 4 Appendix), and manuscript preparation. The thesis supervisor is the second author of this paper.

Acknowledgements

First and foremost I would like to thank my supervisor, Mike Schutz, for his guidance throughout this whole process. I have learned so much over the last few years, and feel quite privileged to work with someone who is incredibly supportive, encouraging and enthusiastic. You've helped me accomplish things I never would have thought possible, and for that I am extremely grateful.

To my committee members, Bruce Milliken and Scott Watter, thank you for your feedback and valuable advice over the last few years. I really appreciate all the time and insights you have given me.

To all the research assistants and students in the lab involved in data collection, data checks and feedback on the manuscripts appearing in this thesis – Aimee Battcock, Lorraine Cheun, Jennifer Harris, Janet Kim, Fiona Manning, Glenn Paul, Olivia Podolak, Matthew Poon, Monique Tardif and Jonathan Vaisberg – Thank-you, thank-you, thank-you, for all of your assistance and support! These projects would not have been the same without the MAPLE Lab team (woot woot!). To the students I had the privileged to supervise – Peter Bamikole, Allison Chan, Kyle Guader, Emily Gula, Monika Kaminski, Grace (Eun Hye) Lee, Zach Louch, Diana Martinez, Kendra Oudyk, Bruce Sheng, Amy Wang – I am so glad I was able to play a part in sparking your interest in research. It was a pleasure to work with each of you.

A big thank you to Jeanine Stefanucci for her work which lead to the inspiration of the project appearing in Chapter 2, and Glenn Paul for his preliminary experiments on the subject. To the reviewers for the manuscript appearing and Chapter 3 – Peter

Pfordresher, Martha Teghtsoonian and Anonymous Reviewer – thank you for your critiques and suggestions. Your feedback has made me strive to improve my writing skills, and become a better scientist and critical thinker.

Lastly, I would like to thank my family and friends for believing in me and supporting me every step of the way. Mom, Dad, Jen, Sal, Sarah & Nathan you have been instrumental pillars in my life, always encouraging me to pursue my dreams. I could not have done this without you.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

GENERAL INTRODUCTION

Amplitude envelope is an important aspect of auditory perception. As one article included (Chapter 3) goes into great detail regarding this, it will not be discussed here. Included are two articles that explore the importance and influence of amplitude envelope in auditory perception research.

The first article (Chapter 2) explores the role of amplitude envelope in an associative memory task, with the aim of improving the associability of auditory alarms in medical devices. Although we found no difference in performance based on amplitude envelope, the paper discusses the patterns of incorrect alarm identification and identifies potential sources of confusion. While this was not our initial goal, we feel this article is a valuable contribution that connects two distinct fields: music cognition and alarm design.

The second article (Chapter 3) encompasses a meta-analysis, surveying the temporal structure of sounds used in auditory perception research, namely in the journal *Attention, Perception & Psychophysics*. This articles discusses several studies in which amplitude envelope has categorically influenced experimental outcomes and suggests that the standard 'flat' temporal structure (i.e. abrupt onset, period of sustain and abrupt offset) may not be the best way to evaluate the auditory system. The goal of this article is to determine what proportion of studies are using the standard 'flat' tones vs. other types of temporal structures we may encounter during everyday listening. These two articles collectively illustrate the original research I have completed on amplitude envelope during my Master's Degree.

CHAPTER 2

Gillard, J. & Schutz, M. (ready for submission). Composing alarms: Considering the musical aspects of auditory alarm design.

Composing alarms: Considering the musical aspects of auditory alarm design

Jessica Gillard
Department of Psychology, Neuroscience & Behaviour, McMaster University, Canada
McMaster Institute for Music and the Mind

Michael Schutz
School of the Arts, McMaster University, Canada
McMaster Institute for Music and the Mind

Corresponding Author: Michael Schutz, School of the Arts, Togo Salmon Hall (TSH), Room 424, McMaster University, 1280 Main St. W, Hamilton, Ontario, L8S 4K1, Canada. Email: schutz@mcmaster.ca

Abstract

Short melodies are commonly linked to referents in jingles, ringtones, movie themes and even auditory displays (i.e. sounds used in human-computer interactions). While melody associations can be quite effective, auditory alarms in medical devices are generally poorly learned and highly confused (Lacherez, Seah, & Sanderson, 2007; Sanderson, Wee, & Lacherez, 2006; Wee & Sanderson, 2008). Here, we draw on approaches and stimuli from both music cognition (melody recognition) and human factors (alarm design) to analyze the patterns of confusions in a paired-associate alarm-learning task involving both a standardized melodic alarm set (Experiment 1) and a set of novel melodies (Experiment 2). Although contour played a role in confusions (consistent with previous research), we observed several cases where melodies with similar contours were rarely confused – melodies holding musically distinctive features. This exploratory work suggests that salient features formed by an alarm's melodic structure (such as repeated notes, distinct contours and easily-recognizable intervals) can increase the likelihood of correct alarm identification. We conclude that the use of musical principles and features may help future efforts to improve the design of auditory alarms.

Introduction

Bidirectional associations between sight and sound are important in many aspects of music. For example, when hearing a familiar piece, some listeners might picture the notation and/or imagine the corresponding movements required for its performance. Likewise, while reading a notated score, musicians will often try to 'hear' the written notes and even envision the correct fingering or movements required for their production. These processes rely in part on associative memory – our ability to make arbitrary cognitive links between cues either within or across modalities.

We make associations involving sound often and with ease in a variety of endeavors, including music, and this skill known as *associative memory* is a well-researched topic. Explorations of word-sound (Godley, Estes, & Fournet, 1984; Keller & Stevens, 2004; Wakefield, Homewood, & Taylor, 2004), image-sound (Bartholomeus & Doehring, 1971; Klingberg & Roland, 1998) and object-sound (Morton-Evans & Hensley, 1978) pairings indicate broad interest in the role of sound in associative memory. However, associative memory studies involving sound are far less frequent than studies of other associations, such as word-word pairings. For example, reviews of word-word studies exploring issues such as concreteness (Paivio, 1971, 1986), structural models (Taylor, Horwitz, & Shah, 2000), and paired associate learning paradigms in the larger study of memory (Roediger, 2008) dominate the literature. The limited focus on sound in associative memory paradigms is surprising, given the importance (and our frequent use) of associations involving sounds in everyday situations.

Sound associations are useful in identifying unseen objects, making appropriate decisions to react (or not) to events around us, and reducing our cognitive load (i.e.

resources put towards different tasks). Sounds are also effective in conveying information, which may be why musical motifs are frequently used in advertising (i.e. jingles) and telecommunications (i.e. ringtones), in order to create cognitive links between sounds and products, corporations or people. Additionally, music plays an important role in movies, operas and plays where associations offer insight into a character's mood (Tan, Spackman, & Bezdek, 2007), or a deeper interpretation of a scene (Vitouch, 2001), such as Wagner's use of leitmotifs or John Williams' use of character themes. Clearly, sound can be an effective medium for conveying information, whether it helps us identify the caller on a phone, announces the arrival of an approaching train, informs us that our email has been sent or foreshadows an important plot development.

Associations in Auditory Alarms

In applied contexts, short musical melodies can serve as the basis for *auditory displays* – sounds used in human-computer interactions. For example, to assist manufacturers the International Electrotechnical Commission (IEC) designed a standardized set of melodic auditory alarms for use in hospitals (i.e. the IEC 60601-1-8 standard).[1] These alarms consist of three- or five-note melodies for medium-priority and high-priority alarms respectively and are used to signal patient-related and machine-related issues to medical practitioners. Unfortunately, problems with the IEC alarms are numerous and well documented. They require extensive exposure to learn (Sanderson et al., 2006; Wee & Sanderson, 2008), are poorly retained (Edworthy & Hellier, 2006; Sanderson et al., 2006) and are frequently confused with one another (Edworthy & Hellier, 2005; Lacherez et al., 2007; Sanderson et al., 2006; Wee & Sanderson, 2008).

However, in these studies participants with at least one year of musical training were better at learning and recalling the IEC Alarms (Sanderson et al., 2006; Wee & Sanderson, 2008).Within each priority level, the IEC alarms have the same length, the same rhythm and span a narrow pitch range (262Hz-523Hz or C4-C5) – characteristics likely contributing to problems in learning and retention (Edworthy, Hellier, Titchener, Naweed, & Roels, 2011; Edworthy & Hellier, 2006; Edworthy & Stanton, 1995; Edworthy, 1994; Sanderson et al., 2006).

The poor discriminability of the IEC alarms has prompted several suggested improvements based on guidelines put forth by alarm deign pioneer, Roy Patterson (Patterson & Mayfield, 1990). These suggested improvements include increasing the heterogeneity of alarms within a set (Edworthy et al., 2011; Phansalkar et al., 2010), varying their contours (Edworthy & Hellier, 2006), and differentiating their rhythms (Edworthy et al., 2011; Edworthy & Hellier, 2006; Edworthy, 2011). However, to the best of our knowledge attempts to apply musical principles to improving their effectiveness have not been widely explored. This is surprising considering that they seem inspired by musical melodies (i.e. most exclusively employ diatonic pitches from a single major scale) and could benefit by research conducted in the field of music cognition.

Associative Memory and Music Cognition

Consulting the music cognition literature, melody recognition and subsequent identification has been suggested to follow the Cohort Theory of spoken word identification (Schulkind, Posner, & Rubin, 2003), originally proposed by Marslen-

Wilson and Tyler (1980). This theory suggests that the initial sound (or notes in the case of music) activate a cohort of possible matches in memory, which is narrowed as the sound (or melody) progresses. Identification occurs once all other candidates are eliminated and a single match is made. Contour plays a fundamental role in melody recognition and recall (Edworthy, 1985; Schulkind et al., 2003), and similarity judgments are largely based on pitch contour, pitch content and inter-onset note patterns (Ahlback, 2007). Together these findings help explain problems with the IEC alarms: Since several of the alarms begin on the same note, have the same general contour, contain many of the same pitches and do not differ in inter-onset note patterns (medium priority alarms are depicted in Figure 1(a)).

Within the music cognition literature, it has been suggested that studies of melody recognition tend to focus on novel vs. familiar stimuli and performance between musicians vs. non-musicians (Müllensiefen & Wiggins, 2011). Consequently, as Schulkind et al. (2003) pointed out, there is little research describing what specific contour patterns actually *facilitate* melody identification. Additionally, Müllensiefen and Wiggins (2011) note that even fewer studies employ paired-associate learning paradigms using melodies. As such, we believe that research combining approaches and stimuli from music cognition (melody recognition) and human factors (alarm design) might offer helpful insights that are relevant to both fields.

Here we describe an exploratory study examining the role of multiple factors that may increase the heterogeneity of alarm design – one common suggestion for improving their efficacy (Edworthy, 2011). One factor of initial interest included amplitude envelope – the shape of a sound over time. Our research team has documented that sounds with natural envelopes (i.e. exponentially decaying 'percussive' sounds) lead to superior performance in an associative memory task over sounds with the artificial sounding flat (i.e. abrupt onset, period of sustain and abrupt offset) envelopes used by the IEC alarms (Schutz et al., 2007). Although amplitude envelope did not appear to play a role in the context of learning and recalling auditory alarms, our exploration did offer useful information regarding the role of melodic structure in confusions. Therefore, the outcomes of these experiments can help inform ongoing efforts to improve the effectiveness of auditory displays by providing insights in the relationship between melodic structure (separate from contour) and confusions.

Experiment 1

In our first experiment, we manipulated the amplitude envelope of the IEC 60601 alarms to be either flat (i.e. the original alarms), or percussive (i.e. the original alarms with reshaped, exponentially decaying envelopes). Based on our team's previous findings, we were interested in investigating whether this parameter might be of use in improving alarm effectiveness. This also allowed us to replicate previous findings regarding patterns of confusion amongst the IEC alarms that differed between experiments using undergraduate students (Sanderson et al., 2006) and experiments using medical professionals (Lacherez et al., 2007; Wee & Sanderson, 2008). This included infrequent confusions of alarms that were phonetically similar (i.e. Perfusion and Power Failure, as well as Perfusion and Infusion) amongst undergraduate students, which in contrast were frequently confused amongst medical professionals in previous studies (Lacherez et al., 2007; Wee & Sanderson, 2008).

Method

Participants. Participants consisted of undergraduate students enrolled in an introductory Psychology course at McMaster University. Forty participants partook in the study for course credit[2] and had on average 3.5 years (SD = 3.7, range = 0-14 years) of musical training.

Stimuli & apparatus. We used the medium priority IEC 60601 alarms and associated referents as stimuli (Figure 1(a)) in a between-subjects design. As the original alarms possess a flat temporal structure, we used the original recordings for our flat condition. To generate percussive versions of the alarms, we reshaped each tone with exponentially decaying envelopes using a MAX/MSP patch previously developed by the MAPLE Lab.[3] We stored the tone sequences on an iMac computer and presented them over Sennheiser HDA 200 headphones at a comfortable listening level held constant for all participants. Prior to beginning the experiment, we administered a short survey including questions regarding demographics, musical training as well as musical practice and listening behaviours.

Procedure. To engage participants, we read them a script asking them to imagine himself or herself as a surgeon and received one of two lists of the eight medical alarm referents (i.e. 'Cardiovascular', 'Perfusion', 'Temperature', etc.), counterbalanced between participants. The experimenter explained the task was to learn to identify eight medical alarms, and defined each of them briefly. The experiment then consisted of four phases: (1) Study Phase, (2) Training Phase, (3) Distracter Task and (4) Evaluation Phase; which are described individually below. We randomized the order in which the alarms were presented for each participant.

Study phase – Participants heard each of the eight alarms twice in a random order along with a verbal statement of the correct alarm referent. We then played a 'masking sound' (white noise through a low-pass filter) for a duration of 6 seconds between different alarms presentations. This blocked out extraneous noise and ensured even spacing between trials to control for individual differences in rehearsal time.

Training phase – We asked participants to identify the correct alarm referent after hearing each of the alarms once in a randomized order. Participants received feedback on their correctness after which we replayed the alarm and reminded them of the correct referent (regardless of their answer). We played the masking sound between sequences once again. Each training block included all eight alarms (played once each). We continued to present training blocks until participants correctly identified 7 out of 8 alarms in two consecutive blocks, or completed a maximum of 10 blocks. To help avoid frustration we offered positive reinforcement every other block (e.g. 'You're doing very well!') regardless of performance.

Distracter task – Upon completion of the training phase, participants performed a silent distracter task (an online mini-golf game[4]) for five minutes.

Evaluation phase – We presented each alarm (randomizing the order for each participant) and asked participants to identify the correct alarm referent. Additionally, we asked participants to indicate how confident they felt about their answer on a scale from 1 (*Not confident at all*) to 6 (*Very confident*). We did not give participants feedback during the Evaluation Phase, but relayed their final score upon completion.

Results

[INSERT FIGURE 1 ABOUT HERE]

Our main manipulation of envelope had no significant affect on performance ($p >$ .05) therefore we collapsed across envelope to analyze the confusion data. On average, participants correctly identified 6.4 (out of a possible 8) alarms in the evaluation phase (SD = 1.57). The patterns of confusion (i.e. when one alarm was 'confused' with another) in the evaluation phase are plotted using the graphics tool Circos[5] in Figure 1(b) (additionally, they are summarized in table form in Appendix A). The plot depicts the total number of confusions (n=62) around the circumference of the circle, with each of the eight alarms represented by different coloured segments according to the following mapping: Oxygen (OX) = Red, Temperature (TE) = Orange, Ventilation (VN) = Olive, General (GE) = Green, Power Failure (PF) = Cyan, Cardiovascular (PE) = Blue, Perfusion (CV) = Purple and Infusion (IN) = Pink.

Longer exterior segments indicate alarms that were highly confused. For example, the long exterior segments for the Ventilation (Olive) and Cardiovascular (Blue) alarms indicate the highest levels of confusion, encompassing 27% (n=17) and 23% (n=14) of total confusions respectively. The medium exterior segments for the Temperature (Orange), Perfusion (Purple) and Infusion (Pink), indicate moderate confusion; accounting for 13% (n=8), 13% (n=8) and 16% (n=10) of total confusions respectively. The relatively short exterior segments for the Oxygen (Red) and General alarms (Green) indicate the least confusion, accounting for only 6.4% (n=4) and 1.6%

(n=1) of total confusions respectively and the Power Failure (Cyan) alarm was not confused at all (n=0). Like-coloured connecting inner bands (i.e. in the same colour as the exterior segments) indicate ‘outbound’ confusions—the times when the alarm in question is confused with another. Different-coloured inner bands indicate ‘inbound’ confusions—the times when other alarms are misidentified as the alarm of interest. The length of each alarm’s exterior segment reflects both its outbound and inbound confusions (smaller segments represent the least-confused alarms). However, as one alarm’s inbound confusions are another’s outbound, we will confine our discussion to the latter to avoid redundancy.

Shorter exterior segments indicate that alarms were confused less frequently. For example, the General alarm (Green) is one of the least confused as indicated by its relatively short exterior segment and thin inner bands of alarm-to-alarm confusion. The like-coloured (i.e. green) inner bands indicate when participants misidentified the General alarm as another, with that band’s thickness reflecting confusion prevalence. For example, only one participant misidentified the General alarm as the Power Failure alarm (Cyan, n=1). The General alarm is one of the least confused (i.e. most successful), for as indicated by its short exterior segment it accounted for only 1.6% (n=1) of the total confusions.

In contrast, the relatively large exterior segment of the Ventilation alarm (Orange) indicates significant confusion. Misidentifications (orange inner bands) include almost all other alarms: Oxygen (Red, n=2), Temperature (Olive, n=3), Power Failure (Cyan, n=1), Cardiovascular (Blue, n=2), Perfusion (Purple, n=5), and Infusion (Pink, n=4), accounting for accounting for 27% (n=17) of total confusions. Additionally, participants frequently misidentified the Ventilation alarm as the Perfusion alarm, indicated by the thick inner orange band spanning across the graph to the Purple (Perfusion) section.

Participants misidentified the Perfusion alarm (Purple) moderately, falling between the extremes of the General and Ventilation alarms, with 13% (n=8) of the total confusions. These included the Oxygen (Red, n=1), Ventilation (Orange, n=6), and Infusion (Pink, n=1) alarms, with the majority stemming from the Ventilation alarm (thick purple inner band extending to the orange Ventilation alarm).

In past investigations, highly confused alarms have been represented by the number of participants that confused one alarm consistently with another on at least 25% of the trials during learning-test cycles (Sanderson et al., 2006; Wee & Sanderson, 2008). This approach is ill-suited for our purposes, since we are only looking at performance during the evaluation phase, and not during the training phase (which is comparable to the learning-test cycles). Therefore, to determine which alarms were 'highly confused', we looked for cells that fell at or above two standard deviations about the mean. In the current data set, any alarm misidentified five or more times consistently with another alarm was considered highly confused (M = 1.1, SD = 1.88). This included confusions between Ventilation and Perfusion (n=5; thick orange band), Cardiovascular and Temperature (n=9; thick blue band), Perfusion and Ventilation (n=6; thick purple band) and, Infusion and Ventilation (n=6; thick Pink band).

Musical training. A t-test revealed that participants with musical training (i.e. one or more years)[6] required significantly fewer training blocks (M=6.4, SD = 2.78) to learn the alarms than participants without musical training (M=8.3, SD = 2.30), t (38) = -2.18, p = .036. However, in the evaluation phase musical training did not significantly affect alarm recall (some training M = 6.7, SD = 1.40; no training M = 5.8, SD = 1.72), t (38) = 1.77, p = .068.

Discussion

We successfully replicated several of the confusions reported in previous studies (Lacherez et al., 2007; Sanderson et al., 2006; Wee & Sanderson, 2008). Many stemmed from similarities in contour, consistent with previous research on contour's role in melody recognition (Edworthy & Hellier, 2006; Edworthy, 1985; Massaro, Kallman, & Kelly, 1980; Schulkind et al., 2003). For example, Temperature (Olive) and Cardiovascular (Blue) both have two ascending intervals; Ventilation (Orange) and Perfusion (Purple) both have an ascending followed by a descending interval. However, we observed a few confusion patterns not explained on the basis of contour.

For example, we observed significant confusions between Ventilation (Orange) and Infusion (Pink) – which differ in contour. A previous study finding similar patterns of confusion attributed this to the fact that the Ventilation and Infusion alarms are often heard together in a medical context (Sanderson et al., 2006). While this has been reported amongst medical professionals (Lacherez et al., 2007; Wee & Sanderson, 2008), this is unlikely to explain our findings here using an undergraduate population lacking exposure to the alarms in medical settings (Sanderson et al. (2006) reported similar findings in a population of students without medical training). This suggests that these confusions might in fact stem from the design of the alarm sequences themselves, rather than the alarm's meaning for medical professionals. We suspect that these alarms might be confused due to 'contour inversion' as opposed to the medical context, as they are essentially mirror images of one another with respect to contour and occupy the same contour space (Dowling, 1971; Marvin & Laprade, 1987).

As with previous studies, we observed low levels of confusion among some alarms. This is consistent with results indicating that distinctive features led to better performance. For example, the Oxygen alarm (Red) is the only melody with two descending intervals in the set and the General alarm (Green) is composed of three repeated notes, making them both easily identifiable and consequently less confusable. Similarly, participants never misidentified the Power Failure alarm (Cyan), consisting of a descending octave followed by a repeated note (i.e. combining two musically salient intervals). These results and insights into the IEC alarm set are consistent with previous observations in both alarm design (Edworthy et al., 2011; Edworthy, 2011) and music cognition (Schulkind et al., 2003).

Despite this consistency, it is also worth mentioning that we failed to replicate several patterns of confusion reported previously amongst medical professionals. For example, the Perfusion and Power Failure alarms were frequently confused by nurses (Lacherez et al., 2007; Wee & Sanderson, 2008), yet they were never confused here. Additionally, one study reported a high prevalence of confusion between the Perfusion and Infusion alarms (Wee & Sanderson, 2008), yet here we found only a mild confusion. These do not appear to be a result of the alarms themselves (i.e. melodies) since they differ in contour and are not frequently confused by undergraduate students reported here and previously (Sanderson et al., 2006), but may rather stem from the alarm referents. We suspect these confusions may be due to contextual cues relevant to medical professionals or even phonetic similarity. Future studies might shed light on this issue by randomizing the IEC alarm sounds and alarm referents.

As suggested by Edworthy, Hellier, Titchener, Naweed and Roels (2011), varying other aspects such as rhythm, timbre and tempo can help reduce misidentification and optimize alarm effectiveness. In other words, increasing heterogeneity among alarms can reduce confusions. Our findings suggest that carefully arranging the pitches according to musical principles can also help to reduce confusion amongst alarms. Similarly, alarms with similar musical characteristics may still be confused even if they differ in contour (i.e. the Ventilation and Infusion alarms).

Experimental design and musical training. Past investigations have suggested that participants with at least one year of musical training are better at learning and recalling the IEC alarms (Sanderson et al., 2006; Wee & Sanderson, 2008). Similarly, we found participants with at least one year of musical training were able to learn the alarms in fewer training blocks compared to participants with no musical training. However here, participants performed equally well on alarm recall in the evaluation phase regardless of whether or not they met the threshold for classification as musically trained (i.e. one year in the alarm literature). In previous studies, participants completed learning-test cycles until they reached 100% accuracy in two consecutive tests or reached a specified time limit ranging from 35 to 50 minutes, and would receive a list of the alarms they identified incorrectly at the end of each test (Lacherez et al., 2007; Sanderson et al., 2006; Wee & Sanderson, 2008). Here, we used a slightly different approach in which participants completed training blocks until they scored at least 7/8 (or 87.5%) in two consecutive blocks or completed a maximum of 10 blocks.

One potentially insightful difference between our design and designs employed previously (showing strong effects of musical background) is that here we replayed the alarm and restated the referent after each response in the training phase, regardless of correctness. This approach may have been particularly helpful to those with less musical exposure, leading to similar performance in the evaluation phase. Future efforts to improve learning and retention of alarms might benefit from exploring these kinds of strategies for those without musical training. Additionally, it suggests that previously reported disadvantages for those without formal musical training may be overcome by changes to the training routine used. Experiment 2 explores this idea, and additionally addresses potential confounds stemming from unvarying melody-referent pairings.

Experiment 2

To further explore whether (a) distinct features help improve correct alarm identification and (b) what aspects appear to group melodies with dissimilar contours as cognitively similar, we looked at confusions in another set of stimuli. Here we paired the same eight IEC alarm referents with eight novel melodies from previous work by our team. This allowed us to determine whether the types of confusions observed with the IEC alarms appear with other melodies. Additionally, we varied the pairings of melodies and alarm referents – an important factor when trying to determine whether confusions stem from melodic structure (i.e. the alarms) or phonetic similarity (i.e. the alarm referents). As previous studies always used set pairings of alarms and referents matching the IEC proposals, this offers a novel chance to disambiguate potential confounds inherent in using the same pairings of sounds and referents for all participants.

Method

Participants. Participants consisted of undergraduate students enrolled in either an introductory Psychology or Linguistics course at McMaster. Forty participants (14 male, 25 female, 1 transgendered) ranging in age from 17 to 24 (M = 19.1 years, SD = 1.56) participated in the study for course credit. Additionally, participants had on average 2.2 years of musical training (SD = 3.07, range = 0-12 years).

Stimuli & apparatus. We selected eight tone sequences consisting of tones drawn from a one octave chromatic scale (A4 – A5) from a set used in a previous study conducted by Schutz and Stefanucci (2010). Each sequence consisted of a sound file with four sine wave (pure tone) notes, roughly 4 seconds in length. Although we manipulated the temporal structure of individual notes within these melodies[7] (as in Experiment 1), here we used a within-subjects design with each participant hearing four melodies with each amplitude envelope. We enumerated the tone sequences (shown in Figure 2(a)) from 1 to 8 and stored them on a MacBook Air laptop. We presented the tone sequences over Sennheiser HDA 200 headphones at a comfortable listening level held constant for all participants. Prior to beginning the experiment, participants completed the survey described in Experiment 1.

Procedure. The procedure for Experiment 2 is similar to that of Experiment 1, with the exception of the use of novel melodies. Additionally, here we randomized the pairings of melodies and alarm referents for all participants rather than maintaining consistent melody-alarm referent pairings, as in the first experiment. In other words, Melody 1 may be paired with the referent 'Ventilation' for one participant, but a paired with a different referent such as ‘Power Failure’ for another participant. This allowed us to control for potentially confusing properties of the referents themselves (i.e. ‘Perfusion’/ ‘Infusion’).

Results

[INSERT FIGURE 2 ABOUT HERE]

Once again, the envelope manipulation had no significant effect on performance ($p > .05$), and we again collapsed across this parameter to analyze the confusion data. Participants correctly identified 5.9 (SD = 1.59) melody-referent pairings in the evaluation phase on average. The patterns of confusion within the evaluation phase are plotted using the graphics tool Circos in Figure 2(b) (and summarized in detail in Appendix B). The plot depicts the total number of confusions (n=85) around the circumference of the circle, with each of the eight melodies represented by different coloured exterior segments according to the following mapping: 1 = Red, 2 = Orange, 3 = Olive, 4 = Green, 5 = Cyan, 6 = Blue, 7= Purple and 8 = Pink. The most highly confused melodies (i.e. the largest exterior segments on the Circos graph) include Melody 3 (Olive), 6 (Blue) and 7 (Purple) representing 22.3% (n=19), 15.3% (n=13) and 21.2% (n=18) of total confusions respectively. Moderately confused melodies include Melody 1 (Red) and 4 (Green) accounting for 12.9% (n=11) and 11.8% (n=10) of total confusions respectively. The least confused included Melodies 2 (Orange), 5 (Cyan) and 8 (Pink), accounting for 5.9% (n=5), 5.9% (n=5) and 4.7% (n=4) respectively. Once again, the thickness of the inner bands between melody segments corresponds to the prevalence of their confusion.

Again, to determine which alarms were highly confused, we looked for cells that fell at or above two standard deviations about the mean. Here, any melody misidentified five or more times consistently as another alarm was considered highly confused (M = 1.5, SD = 1.77). This included confusions between Melody 3 and Melody 6 (n=10; thick olive band), as well as confusions between Melody 7 and Melodies 3 and 4 (n=5 each; thick purple bands).

Musical training. In contrast to the previous experiment, participants with at least one year of musical training did not learn the melody-referent associations any faster (M = 7.2, SD = 2.387) than participants without musical training (M = 8.2, SD = 2.34), $t(38) = -1.38$, $p = .175$. Additionally, their performance in the evaluation phase did not differ significantly (some training M = 6.2, SD = 1.58; no training M = 5.6, SD = 1.57), $t(38) = 1.281$, $p = .208$. Furthermore in comparison to Experiment 1, participants in Experiment 2 had significantly fewer years of musical training (Experiment 1: M = 3.5, SD = 3.70; Experiment 2: M = 2.2, SD = 3.03), $t(39.597) = 2.19$, $p = .034$.

Discussion

Overall, these melodies yielded greater confusions (85) than the alarms used in Experiment 1 (62). This may be attributed to their increased length (4 notes rather than 3) and the greater variety of intervals used. In Experiment 1, the IEC alarms consisted of 3-note melodies, composed almost entirely from the Major scale. In Experiment 2, the to-be-learned melodies consisted of 4 notes that are less strictly diatonic in their structure.

Previous research has shown a decrease in contour judgment performance with increasing melody length (Edworthy, 1982, 1985) as well as poorer recognition of atonal vs. tonal melodies (Mikumo, 1992).

Again, it appears that contour plays an important role in explaining this pattern of confusions. The most highly confused melodies (3 and 6) share a contour consisting of an ascending interval followed by two descending intervals. Interestingly, the final interval in both melodies is a descending major second. Several participants (25%) misidentified Melody 3 as Melody 6, which accounted for over *half* of its total confusions (53%). However, once again there are systematic results that are not explained by contour.

Curiously, other melody pairs with similar contours—arranged vertically in Figure 2(a) (i.e. 1 with 8, 2 with 7, and 4 with 5)— were *not* frequently confused. Moreover, we found high rates of confusion between alarms differing substantially in contour. As in the first experiment, we suspect this reflects the importance of musically distinctive features (i.e. a repeated note in Melody 1 vs. Melody 8, prominent octave interval in Melody 2 vs. Melody 6, and a salient change in contour in Melody 4 vs. Melody 5).

For example Melody 7 contains a descending interval followed by an ascending and then descending interval (Figure 2(a)). Participants confused this melody with Melodies 1, 2, 3, 4 and 6. Despite their dissimilarities, the contours of all five melodies contain changes of direction. It is possible that melodies that contain one, or multiple changes in direction with no other defining features are more easily confused. This might also explain why participants confused Melody 7 less often with Melody 2. Although these melodies are very similar in contour, the repeated note in Melody 2 may have acted as a distinct feature allowing participants to better differentiate the two. Additionally, the descending octave – an interval that is very salient even to an untrained ear (Krumhansl & Kessler, 1982) – may have helped differentiate Melody 2.

Of the melodies that contain changes in direction, Melody 1 and 4 seem to be less confused than their counterparts, despite having one change of pitch direction each. This may be due to more subtle yet still distinct features. For example, Melody 1 contains a tritone, likely making it sound sadder (Huron, 2008) and less stable (Krumhansl & Kessler, 1982) than the other melodies. Some participants, particularly those with significant musically training, may have been able to identify this and use it in their learning. Likewise, Melody 4 consists of two descending intervals followed by an ascending interval, beginning and ending on the same note. This return back to the initial note, or the 'tonic' of these 4 note melodies, has been shown to improve melody recognition (Krumhansl, 1979) and may have helped differentiate Melodies 4 and 5, which have exactly the same contour with the exception of the last interval.

These salient intervals may help minimize confusion. The same can be said for Melodies 5 and 8 in that distinctive features, such as an overall descending contour (as in Melody 5), or a successively repeated note (as in Melody 8) are highly salient and can easily be differentiated from other alarms.

Melody-referent confusions. In this experiment we randomized the melodies and alarm-referents, allowing us to address confusions caused by the referents themselves (i.e. phonetic similarity). As with previous studies, we saw modest confusions between

Perfusion and Power Failure as well and Perfusion and Infusion referents. This suggests potential problems with the alarm names independent of the alarm sequences – a possibility that to the best of our knowledge has not been reported, as most studies tend to associate the alarms only with their recommended referent commands, confounding interpretations of confusions. Additionally, if phonetically similar alarm referents are paired with melodies that are also very similar, confusions could be additive and subsequently compound the problem.

Musical training. In Experiment 1 we found that musically trained participants required significantly fewer training blocks than musically untrained participants. Yet here we found no such effect, with musically trained and untrained participants performing similarly in both the training and evaluation phases. This may be attributed to the fact that participants in Experiment 2 had significantly fewer years of musical training overall than participants in Experiment 1. Consequently, musically trained participants took on average 6.4 blocks in Experiment 1 to learn the association, but 7.2 blocks in Experiment 2 (musically untrained participants did not differ between the two experiments – requiring 8.3 blocks in the first and 8.2 in the second experiment). It is possible that a certain level of training is required to affect performance on this task.

Conclusion

IEC Alarm Confusions

Ensuring alarm sets are efficient and memorable is a significant and timely issue in human factors and alarm design, the subject of intensive studies offering a plethora of ideas for improvements (Edworthy et al., 2011; Edworthy, 2011; Phansalkar et al., 2010;

Sanderson, Liu, & Jenkins, 2009). However, these studies rarely focus on the musical structure of auditory alarms, such as how particular mixes of musical intervals contribute to confusions. This is somewhat surprising, given that melody recognition is a rich topic within the field of music perception. To contribute towards efforts bridging these areas of research, here we explored alarm learning using both a standard alarm set (i.e. the IEC alarms) in Experiment 1 and a novel alarm set in Experiment 2. By randomizing the melody-referent pairings in Experiment 2, we were also able to avoid potential confounds inherent when using the same alarm-referent pairs (unavoidable in previous experiments for obvious reasons).

Additionally, our results suggest that the superior performance of musically trained participants (i.e. having at least one year of formal musical training) reported in previous studies may be attributed to the training structure of the task. Unlike previous investigations where participants received a list of alarms they identified incorrectly (Lacherez et al., 2007; Sanderson et al., 2006; Wee & Sanderson, 2008), we reinforced the alarm/melody and referent after every trial, regardless of response correctness. Furthermore, we believe this directly affected performance in the evaluation phases, where we found no significant difference in performance between participants with some musical training vs. no musical training.

Although more research is required to fully explore the ideas raised by our findings, they suggest that the accuracy of identifying melodic alarms may be improved by varying the tonal qualities of alarms and including salient features such as repeated notes, distinct contours and distinctive intervals (i.e. by avoiding focusing exclusively on notes from within a major scale, which can limit opportunities for heterogeneity).

Consequently, attention to the melodic structure of auditory alarms offers another technique for increasing heterogeneity in alarm sets—a factor relevant to on-going efforts to improve alarm design. These principles could be used to build more robust redundancy into alarm cues by co-varying interval quality and tone durations, for example.

It is important to note that even small improvements in auditory alarm design could lead to potentially large improvements in patient care. For example, one observational study in an Intensive Care Unit (ICU) found that on average, two critical alarms are missed per day *per hospital* (Donchin et al., 2003). Consequently, in the United States with approximately 5,720 registered hospitals servicing a population of 315 million, this rate corresponds to roughly 4.2 million errors per year. We recognize that medical professionals respond to many more alarms than they miss, and also that there are problematic aspects of alarms beyond their structure. However, we mention this issue here to underscore both the magnitude of the problem as well as the powerful potential public health benefits of even incremental improvements in alarm design by any means – such as through attention to basic principles of melodic structure.

Broader Implications for Music Cognition Research

While it is clear that we are able to easily make associations with music (as in the case of jingles, ringtones and musical motifs in operas, plays or movies), it is less clear which specific features facilitate (or hinder) melody identification and the subsequent retrieval of these associations. Our exploratory data provides some insight on this issue; suggesting that distinctive features (i.e. repeated notes, distinctive contours, variations in tonality, etc.) are important factors in helping to distinguish melodies. Although future research is needed to further test some of the ideas arising from our data, we believe this work holds value in improving our understanding of associative memory involving sounds, as well as informing research on melody identification. This also provides a unique opportunity in which music cognition research may be used in an applied setting to inspire future efforts to improve the design of auditory alarms – an issue of broad relevance to public health.

Acknowledgements

We would like to thank Janet Kim, Fiona Manning, Glenn Paul, Olivia Podolak, Matthew Poon and Jonathan Vaisberg for their assistance in data collection, and Jeanine Stefanucci for her assistance in exploring the ideas leading to this project. We would like to acknowledge financial assistance for this research through grants to Dr. Michael Schutz from the Natural Sciences and Engineering Research Council of Canada (NSERC RGPIN/386603-2010), Ontario Early Researcher Award (ER10-07-195), McMaster University Arts Research Board, and the Canadian Foundation for Innovation (CFI-LOF 30101).

Notes

1. http://itee.uq.edu.au/%7Ecerg/auditory/alarms.htm
2. Demographic information could not be provided due to an unfortunate lab flooding in which we lost all hardcopies of participant information collected for this experiment before it could be saved electronically.

3. Flat alarms consisted of three tones 244ms in length (including 25ms rise/fall times) separated by 156ms. The general structure of percussive alarms was the same with the exception of the individual tones having a rise time of 25ms, followed by an immediate exponential decay for the remaining duration of the tone.
4. http://www.addictinggames.com/sports-games/miniputt3.jsp
5. http://circos.ca/circos_online/
6. Previous explorations of IEC alarm learning classified individuals with least one year of musical training as 'musically trained.'
7. We used SuperCollider to shape pure tones (i.e. sine waves) into flat and percussive envelopes to create individual tones. We then arranged these individual into sequences using Audacity – a free sound-editing program. All tone sequences consisted of four one-second sound clips, either all percussive or all flat, concatenated together to create a four-second melody. Percussive tones were approximately 800ms in length separated by approximately 150ms. Flat tones were 745ms in length separated by 200ms.

References

Ahlback, S. (2007). Melodic similarity as a determinant of melody structure. *Musicae Scientiae*, *11*(1), 235–280.

Bartholomeus, B., & Doehring, D. (1971). Effect of stimulus order on verbal and nonverbal visual-auditory association. *Perceptual and Motor Skills*, *33*(3), 891–897.

Donchin, Y., Gopher, D., Olin, M., Badihi, Y., Biesky, M., Sprung, C. L., … Cotev, S. (2003). A look into the nature and causes of human errors in the intensive care unit. 1995. *Quality & Safety in Health Care*, *12*(2), 143–148.

Dowling, W. (1971). Recognition of inversions of melodies and melodic contours. *Perception & Psychophysics*, *9*(3B), 348–349.

Edworthy, J. (1982). Pitch and contour in music processing. *Psychomusicology: A Journal of Research in Music Cognition*, *2*(1), 44–46.

Edworthy, J. (1985). Interval and contour in melody processing. *Music Perception*, *2*(3), 375–388.

Edworthy, J. (1994). The design and implementation of non-verbal auditory warnings. *Applied Ergonomics*, *25*(4), 202–210.

Edworthy, J. (2011). Designing Effective Alarm Sounds. *Biomedical Instrumentation & Technology*, (August), 290–294.

Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *Quality & Safety in Health Care*, *14*(3), 212–215.

Edworthy, J., & Hellier, E. (2006). Alarms and human behaviour: implications for medical alarms. *British Journal of Anaesthesia*, *97*(1), 12–17.

Edworthy, J., Hellier, E., Titchener, K., Naweed, A., & Roels, R. (2011). Heterogeneity in auditory alarm sets makes them easier to learn. *International Journal of Industrial Ergonomics*, *41*(2), 136–146.

Edworthy, J., & Stanton, N. (1995). A user-centred approach to the design and evaluation of auditory warning signals: 1. Methodology. *Ergonomics*, *38*(11), 2262–2280.

Godley, R., Estes, R., & Fournet, G. (1984). Paired-associate learning as a function of age and mode of presentation. *Perceptual and Motor Skills*, *59*(3), 959–965.

Huron, D. (2008). A Comparison of Average Pitch Height and Interval Size in Major- and Minor-key Themes : Evidence Consistent with Affect-related Pitch Prosody. *Empirical Musicology Review*, *3*, 59–63.

Keller, P., & Stevens, C. J. (2004). Meaning from environmental sounds: types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied*, *10*(1), 3–12.

Klingberg, T., & Roland, P. E. (1998). Right prefrontal activation during encoding, but not during retrieval, in a non-verbal paired-associates task. *Cerebral Cortex*, *8*(1), 73–79.

Krumhansl, C. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, *11*(3), 346–374.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368.

Lacherez, P., Seah, E. L., & Sanderson, P. M. (2007). Overlapping Melodic Alarms Are Almost Indiscriminable. *Human Factors*, *49*(4), 637–645.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*(1), 1–71.

Marvin, E. W., & Laprade, P. a. (1987). Relating Musical Contours: Extensions of a Theory for Contour. *Journal of Music Theory*, *31*(2), 225–267.

Massaro, D. W., Kallman, H. J., & Kelly, J. L. (1980). The role of tone height, melodic contour, and tone chroma in melody recognition. *Journal of Experimental Psychology. Human Learning and Memory*, *6*(1), 77–90.

Mikumo, M. (1992). Encoding Strategies for Tonal and Atonal Melodies. *Music Perception*, *10*(1), 73–82.

Morton-Evans, A., & Hensley, R. (1978). Paired associate learning in early infantile autism and receptive developmental aphasia. *Journal of Autism and Childhood Schizophrenia*, *8*(1), 61–69.

Müllensiefen, D., & Wiggins, G. (2011). Sloboda and Parker 's recall paradigm for melodic memory : A new, computational perspective. In *Essays in Honour of John Sloboda* (pp. 161–186).

Paivio, A. (1971). *Imagery and Verbal Processes*. New York, NY: Holt, Rinehart & Winston.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, England: Oxford University Press.

Patterson, R. D., & Mayfield, T., F. (1990). Audiotry warning sounds in the work environment [and discussion]. *Phiolspohical Transactions of the Royal Society of London*, *327*(1241), 485–492.

Phansalkar, S., Edworthy, J., Hellier, E., Seger, D. L., Schedlbauer, A., Avery, A. J., & Bates, D. W. (2010). A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems. *Journal of the American Medical Informatics Association*, *17*(5), 493–501.

Roediger, H. L. (2008). Relativity of remembering: why the laws of memory vanished. *Annual review of psychology*, *59*, 225–54.

Sanderson, P. M., Liu, D., & Jenkins, S. A. (2009). Auditory displays in anesthesiology. *Current Opinion in Anaesthesiology*, *22*(6), 788–795.

Sanderson, P. M., Wee, A. N., & Lacherez, P. (2006). Learnability and discriminability of melodic medical equipment alarms. *Anaesthesia*, *61*(2), 142–147.

Schulkind, M. D., Posner, R. J., & Rubin, D. C. (2003). Musical features that facilitate melody identification: How do you know it's "your" song when they finally play it? *Music Perception*, *21*(2), 217–249.

Schutz, M., & Stefanucci, J. K. (2010). Amplitude Envelope and Auditory Alarms. In *Proceedings of the International Conference of Music Perception and Cognition*. Seattle, WA.

Schutz, M., Stefanucci, J. K., Carberry, A., & Roth, A. (2007). Name that (Percussive) Tune: Tone Envelope Affects Learning. In *Proceedings of the Auditory Perception, Cognition and Action Meeting Conference*. Long Beach, CA.

Tan, S., Spackman, M., & Bezdek, M. (2007). Viewers' interpretations of film characters' emotions: Effects of presenting film music before or after a character is shown. *Music Perception*, *25*(2), 135–152.

Taylor, J., Horwitz, B., & Shah, N. (2000). Decomposing memory: functional assignments and brain traffic in paired word associate learning. *Neural Networks*, *13*, 923–940.

Vitouch, O. (2001). When Your Ear Sets the Stage: Musical Context Effects in Film Perception. *Psychology of Music*, *29*(1), 70–83.

Wakefield, C. E., Homewood, J., & Taylor, A. J. (2004). Cognitive compensations for blindness in children: An investigation using odour naming. *Perception*, *33*(4), 429–442.

Wee, A. N., & Sanderson, P. M. (2008). Are melodic medical equipment alarms easily learned? *Anesthesia & Analgesia*, *106*(2), 501–508.

Figure 1. Contours of the eight IEC 60601 alarms (a) and Confusions in the Evaluation Phase of Experiment 1 (b). Alarms are represented by colours: Oxygen (OX) = Red, Ventilation (VN) = Orange, Temperature (TE) = Olive, General (GE) = Green, Power Failure (PF) = Cyan, Cardiovascular (CV) = Blue, Perfusion (PE) = Purple and Infusion (IN) = Pink. In panel (a), M = Major, m = Minor, P = Perfect, TT = Tritone. + = Ascending, - = Descending. In panel (b), thicker exterior segments and inner bands (connecting two segments) indicate higher rates of confusion. Inner bands in the same colour as the exterior segment indicate the times when the alarm in question is confused with another (i.e. outbound confusions). Inner bands of different colours than the exterior segment indicate the times when other alarms are confused with the alarm represented by the exterior segment (i.e. inbound confusions).

Figure 1.

(a)

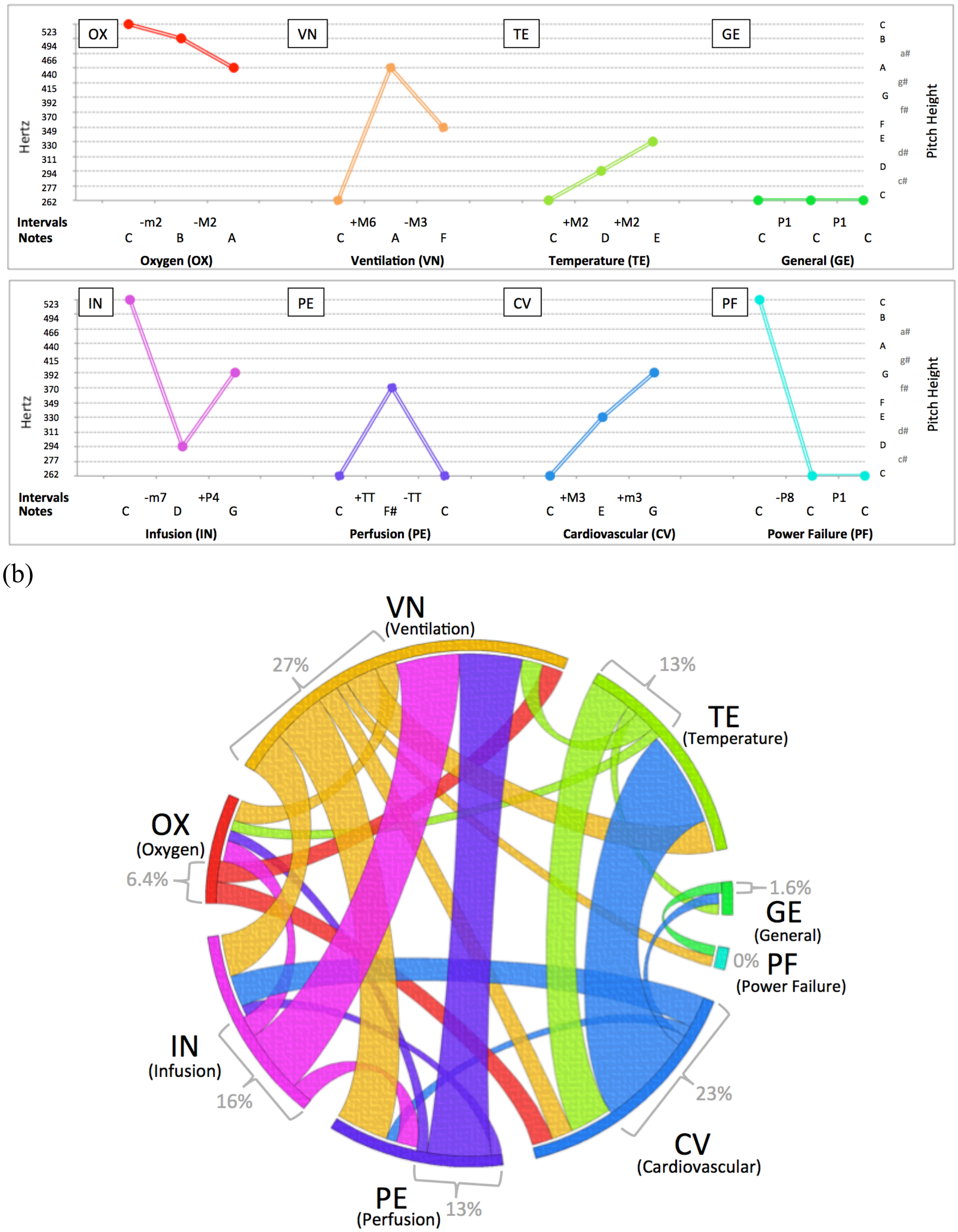
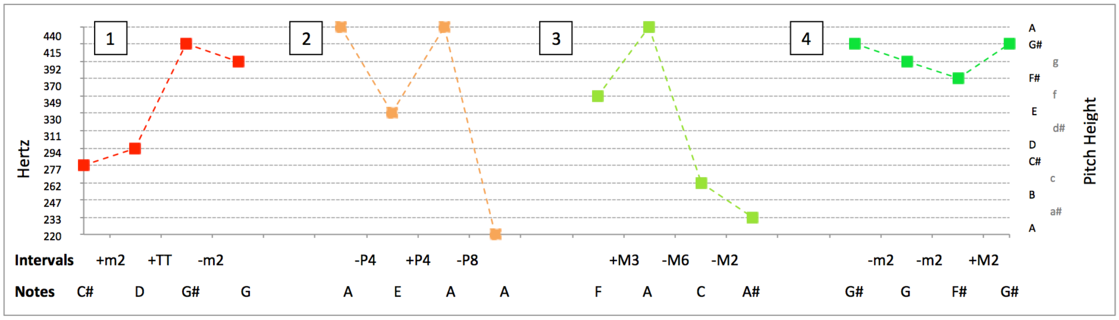

OX
VN
TE
GE
Hertz
523
494
466
440
415
392
370
349
330
311
294
277
262
Pitch Height
Intervals
Notes
-m2 -M2
C B A
Oxygen (OX)
+M6 -M3
C A F
Ventilation (VN)
+M2 +M2
C D E
Temperature (TE)
P1 P1
C C C
General (GE)
IN
PE
CV
PF
-m7 +P4
C D G
Infusion (IN)
+TT -TT
C F# C
Perfusion (PE)
+M3 +m3
C E G
Cardiovascular (CV)
-P8 P1
C C C
Power Failure (PF)
(b)
VN
(Ventilation)
27%
13%
TE
(Temperature)
OX
(Oxygen)
6.4%
1.6%
GE
(General)
0%
PF
(Power Failure)
IN
(Infusion)
16%
23%
CV
(Cardiovascular)
PE
(Perfusion)
13%

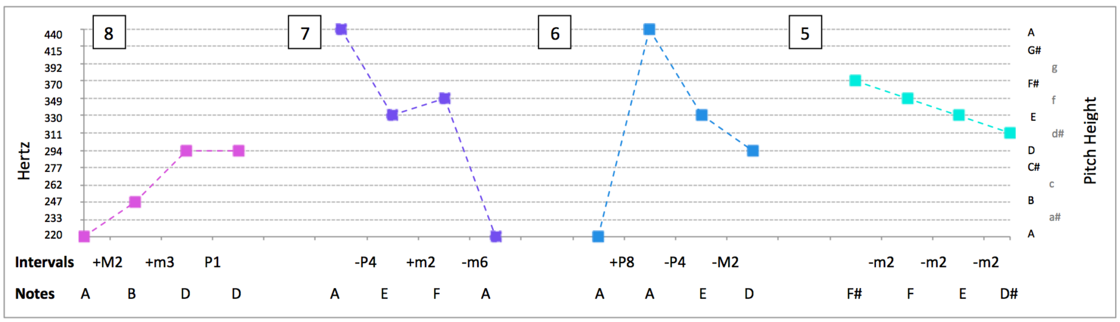
Figure 2. Contours of the eight melodies (a) and Confusions in the Evaluation Phase of Experiment 2 (b). In panel (a), M = Major, m = Minor, P = Perfect, TT = Tritone. + = Ascending, - = Descending. In panel (b), melodies are distinguished by colour: 1 = Red, 2 = Orange, 3 = Olive, 4 = Green, 5 = Cyan, 6 = Blue, 7= Purple and 8 = Pink. Thicker exterior segments and inner bands indicate higher rates of confusion. Inner bands in the same colour as the exterior segment indicate the times when the melody in question is confused with another (i.e. outbound confusions). Inner bands of different colours than the exterior segment indicate the times when other melodies are confused with the alarm represented by the exterior segment (i.e. inbound confusions). Colours are for clarifying individual alarms/melodies within an experiment. There is no intended correspondence between colors within the two experiments (i.e. Oxygen in Exp. 1 and alarm 1 in Exp. 2 are both Red but are not necessarily related in any way.

Figure 2.

(a)

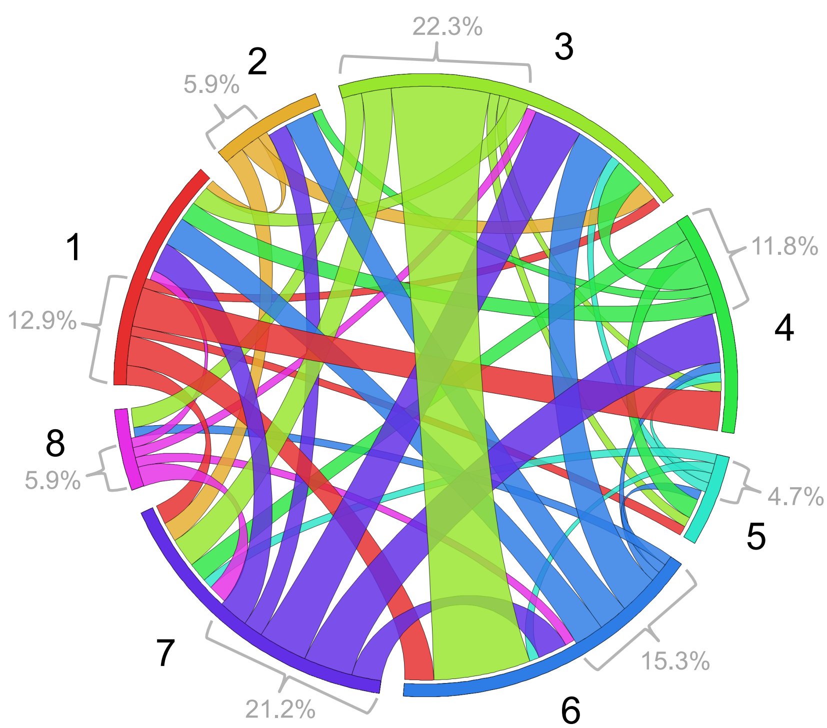


1
2
3
4
Hertz
440
415
392
370
349
330
311
294
277
262
247
233
220
Pitch Height
A
G#
g
F#
f
E
d#
D
C#
c
B
a#
A
Intervals +m2 +TT -m2 -P4 +P4 -P8 +M3 -M6 -M2 -m2 -m2 +M2
Notes C# D G# G A E A A F A C A# G# G F# G#



8
7
6
5
Hertz
Pitch Height
Intervals +M2 +m3 P1 -P4 +m2 -m6 +P8 -P4 -M2 -m2 -m2 -m2
Notes A B D D A E F A A A E D F# F E D#

(b)



1
2
3
4
5
6
7
8
22.3%
5.9%
11.8%
12.9%
5.9%
4.7%
15.3%
21.2%

Appendix A

Summary of Experiment 1 Alarm Confusions

	Oxygen (Red)	Ventilation (Orange)	Temperature (Olive)	General (Green)	Power Failure (Cyan)	Cardiovascular (Blue)	Perfusion (Purple)	Infusion (Pink)
OX		n=2	n=1	n=0	n=0	n=0	n=1	n=2
VN	n=2		n=2	n=0	n=0	n=0	n=6	n=6
TE	n=0	n=3		n=0	n=0	n=9	n=0	n=0
GE	n=0	n=0	n=1		n=0	n=1	n=0	n=0
PF	n=0	n=1	n=0	n=1		n=0	n=0	n=0
CV	n=2	n=2	n=4	n=0	n=0		n=0	n=0
PE	n=0	n=5	n=0	n=0	n=0	n=1		n=2
IN	n=0	n=4	n=0	n=0	n=0	n=3	n=1	
	Total = 4 (6.4%)	**Total = 17 (27%)**	**Total = 8 (13%)**	**Total = 1 (1.6%)**	**Total = 0 (0%)**	**Total = 14 (23%)**	**Total = 8 (13%)**	**Total = 10 (16%)**

Note. Total confusions for Experiment 1, n= 62

Appendix B

Summary of Experiment 2 Melody Confusions

	1 (Red)	2 (Orange)	3 (Olive)	4 (Green)	5 (Cyan)	6 (Blue)	7 (Purple)	8 (Pink)
1		n=1	n=2	n=2	n=0	n=3	n=3	n=1
2	n=0		n=0	n=1	n=0	n=3	n=2	n=0
3	n=1	n=2		n=3	n=1	n=4	n=5	n=1
4	n=4	n=0	n=1		n=1	n=1	n=5	n=0
5	n=1	n=0	n=1	n=2		n=1	n=0	n=0
6	n=3	n=0	n=10	n=0	n=1		n=3	n=1
7	n=2	n=2	n=3	n=2	n=1	n=0		n=2
8	n=0	n=0	n=2	n=0	n=0	n=1	n=0	
	Total = 11 (12.9%)	**Total = 5 (5.9%)**	**Total = 19 (22.3%)**	**Total = 10 (11.8%)**	**Total = 4 (4.7%)**	**Total = 13 (15.3%)**	**Total = 18 (21.2%)**	**Total = 5 (5.9%)**

Note. Total confusions for Experiment 2, n= 85

CHAPTER 3

Gillard, J. & Schutz, M. (ready for submission). Classifying the properties of sounds used in auditory perception research.

Classifying the properties of sounds used in auditory perception research

Jessica Gillard[1,3], Michael Schutz[2,3,*]

[1]Department of Psychology, Neuroscience & Behaviour, McMaster University, Canada

[2]School of the Arts, McMaster University, Canada

[3]McMaster Institute for Music and the Mind

[*]Corresponding Author: Michael Schutz, School of the Arts, Togo Salmon Hall (TSH), Room 424, McMaster University, 1280 Main St. W, Hamilton, Ontario, L8S 4K1, Canada. Telephone: +1-905-525-9140 Ext. 23159. Email: schutz@mcmaster.ca

Abstract

Many of the sounds we hear possess dynamic temporal structures rich in information. Researchers have previously speculated that auditory experiments disproportionately employ stimuli with simplistic temporal structures (Neuhoff, 2004). This raises questions regarding whether their conclusions generalize to the processing of sounds with dynamic amplitude changes – a common characteristic of natural sounds. To explore the issue empirically, we conducted a novel survey of *Attention, Perception & Psychophysics* to establish a baseline understanding of the sounds used in auditory research. A detailed analysis of 210 experiments from 94 articles selected evenly from the journal's history reveals that 93% of stimuli employed temporal structures lacking the dynamic variations characteristic of sounds heard outside the laboratory. Given differences in task outcomes and even the underlying perceptual strategies evoked by dynamic vs. invariant temporal structures, this heavy focus on one type of stimuli raises important questions of broad relevance. As this survey was based on a representative sample of publications in a prominent journal, these results suggest that stimuli with time-varying temporal shapes offer significant potential for furthering our understanding of the perceptual system's structure and function.

Introduction

Sounds synthesized with temporal shapes (aka "amplitude envelopes") consisting of rapid onsets followed by sustain periods and rapid offsets afford precise quantification and description – qualities of obvious methodological value. However as William Gaver argued in a different context, fixating on simplistic sounds can obfuscate the processes used in everyday listening situations (Gaver, 1993a, 1993b). A sound's temporal structure is rich in information, informing listeners about the materials involved in an event (Klatzky, Pai, & Krotkov, 2000; Lutfi, 2007), or even an event's outcome – such as whether a dropped bottle bounced or broke (Warren & Verbrugge, 1984). The simplistic structures of "flat" tones lack such dynamic temporal information.

Members of our research team have documented markedly different task outcomes when using tones with simplified/invariant vs. natural/dynamic temporal structures on tasks ranging from audio-visual integration (Schutz, 2009), to associative memory (Schutz, Stefanucci, Carberry, & Roth, under review) and even underlying processing strategies (Vallet, Shore, & Schutz, in press). Examples of flat and percussive tones used in these experiments can be seen in Figure 1. These ongoing projects complement previous work documenting perceptual differences in the processing of tones with "ramped" or "looming" (i.e. increasing in intensity over time) vs. "damped" or "receding" (i.e. decreasing in intensity over time) temporal structures. Although time-reversed but otherwise identical, these spectrally matched sounds are perceived as differing in duration (DiGiovanni & Schlauch, 2007; Grassi & Darwin, 2006; Grassi & Pavan, 2012; Grassi, 2010; Schlauch, Ries, & DiGiovanni, 2001), loudness (Ries, Schlauch, & DiGiovanni, 2008; Stecker & Hafter, 2000; Teghtsoonian, Teghtsoonian, &

Canévet, 2005), and loudness change (Neuhoff, 1998, 2001). Additionally, they differentially integrate with visual information in both perceptual organization (Grassi & Casco, 2009) and duration judgment (Schutz, 2009) tasks.

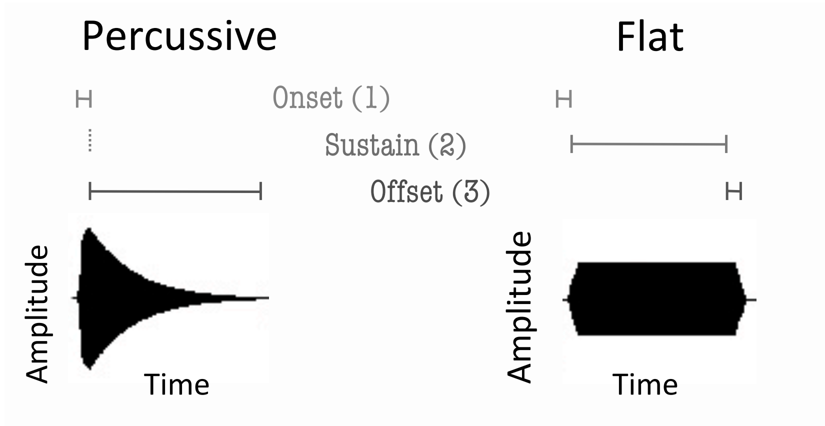


Figure 1. Examples of percussive (left) and flat (right) tones. After onset, percussive tones immediately begin exponential decays whereas flat tones exhibit sustain periods of indefinite length followed by abrupt offsets.

Temporal Structure's Important Role in Perceptual Processing

Experiments conducted primarily with flat tones might suggest conclusions that do not generalize to natural sounds. For example, although vision is generally thought to have minimal influence on auditory judgments of event duration (Guttman, Gilroy, & Blake, 2005; Walker & Scott, 1981; Welch & Warren, 1980), it strongly affects judgments of musical note duration made when watching videos of a professional percussionist making long and short gestures (Schutz & Lipscomb, 2007). This robust effect replicates when using impact (but not sustained) sounds from other events (Schutz & Kubovy, 2009a), point-light simplifications of the visual information (Schutz & Kubovy, 2009b) and even a single moving dot (Armontrout, Schutz, & Kubovy, 2009). The exceptional nature of this integration stems from the dynamically decaying temporal structure of natural, impact sounds. Simplifications of the auditory component using pure tones shaped with temporal structures characteristic of impacts integrate with visual information, whereas pure tones shaped with amplitude invariant, flat, structures do not. The latter finding is consistent with previous research demonstrating vision's lack of influence on auditory judgments of event duration. As such, theories derived from experiments using temporally simplistic stimuli fail to generalize to sounds with the kinds of dynamic structures found in impact events (Schutz, 2009).

Temporal structure's effect on audio-visual integration is not limited to duration perception. Two disks moving across a screen and briefly crossing paths generally appear to 'pass through' one another (Sekuler, Sekuler, & Lau, 1997), and a click simultaneous with the overlap increases the probability of instead seeing a 'bounce.' However, subsequent research has found that damped (i.e. decreasing in intensity over time) sounds elicit stronger bounce percepts than ramped (i.e. increasing in intensity over time) sounds, presumably as they are event-consistent (Grassi & Casco, 2009).

Temporal Structure and Auditory Perception

In addition to affecting perceived duration, temporal structure can affect the *underlying strategies* used in auditory processing. The durations of amplitude invariant tones can be evaluated using a 'marker strategy' – marking tone onset and offset, consistent with Scalar Expectancy Theory (Gibbon, 1977; Machado & Keen, 1999). However, this strategy would be ill-suited for dynamic sounds with decaying offsets, as their moment of acoustic completion is ambiguous. Recent work suggests such sounds' durations (which are common in our environment, Gaver, 1993a, 1993b), can be processed with a ‘prediction strategy’ estimating tone completion from decay rate (Vallet et al., in press), explaining why duration judgments are no more variable for gradually decaying vs. abruptly ending tones (Schutz, 2009). Because temporal structure affects performance on a variety of tasks, understanding its use in auditory research is of broad interest.

What Sounds are used in Auditory Perception Research?

Although simplistic temporal structures may be sensible for certain individual experiments, relying disproportionately on them could lead to problematic conclusions about the auditory system’s structure, which evolved in a world filled with sounds exhibiting dynamic amplitude changes (see Figure 2c for samples of natural sounds). We note a parallel in this respect with visual perception, where the crucial role of a stimuli’s three-dimensional structure (Snow et al., 2011) and movement of both stimulus and observer (Gibson, 1954) go undetected in research employing only static 2D images. Understanding whether temporal structure’s role may be similarly underappreciated requires knowledge of its use in auditory perception research. Here we contribute to this goal by analyzing a representative sample of experiments drawn from several decades of publications in *Attention, Perception & Psychophysics.*

A previous survey of *Music Perception* revealed surprisingly that 35% of studies omitted definition of their stimuli’s temporal structure (Schutz & Vaisberg, 2014). However, that survey included only relatively simple experiments using single tones or isolated series of tones, and focused heavily on music. Furthermore, it drew unequally from different time periods, making it difficult to discern broad trends. In order to assess a true representative sample, we conducted a survey (a) exploring auditory perception broadly beyond music, (b) incorporating diverse paradigms with many different experimental designs and stimuli, (c) tracking a variety of properties for all stimuli (i.e. temporal structure, duration, etc.) to contextualize relative levels of detail, and (d) of a journal widely recognized for its rigor. Consequently, this survey lends useful insight into whether auditory perception research relies heavily on stimuli compelling conclusions that do not fully explain the perceptual system's abilities and behavior.

Method

In order to obtain a broad sample of auditory experiments, we searched the database *PsycInfo* using the terms 'Perception & Psychophysics' (Publication title), 'Auditory' (identifier/keyword), and NOT 'speech', 'language', 'phonetic' and 'dialect' to identify auditory studies of non-speech sounds. We selected the first two articles within each year (i.e. 1966 - 2012) of the resulting 422, yielding a corpus of 94 papers. This stratified quota sampling technique is consistent with best practices for accurate sampling in public opinion polls and market research (Smith, 1983), and the inclusion of 22% of the 422 articles exceeds common sampling benchmarks (Jackman, 2005; Scheuren, 2004),

We coded all experiments (n=210) individually within the 94 articles, classifying only the auditory components of multisensory stimuli. Due to the complexity of these experiments, we distributed one point amongst all sound categories within each experiment, rather than merely between all sounds used as done in our team's recent survey of *Music Perception* (Schutz & Vaisberg, 2014). For example, if an experiment used two sound categories (i.e. a target and distractor), each sound category received a half point. In an experiment with four types of targets and two types of distractors, each target and distractor received 0.125 and 0.25 points respectively (sample point weightings appear in Table 1).

Table 1

Examples of Point Weighting Distributions

Article	Experiment Number	Sound Categories	Functional Category	Point Weighting	Envelope Category
Radeau & Bertelson, 1978	1	1	stimulus	1.00	percussive
Shinn-Cunningham, 2000	1	1	target	1.00	click train
Five experiments, each using a single type of sound	2		target	1.00	click train
	3		target	1.00	click train
	4		target	1.00	click train
	5		target	1.00	click train
Boltz, Marshburn, Jones, & Johnson, 1985.	1	2	stimulus	0.50	flat
	1		warning tone	0.50	undefined
Two experiments, each using two types of sounds	2	2	stimulus	0.50	flat
	2		warning tone	0.50	undefined
Stilp, Alexander, Keifte & Kluender, 2010	1	3	target A	0.333	other
	1		target B	0.333	other
Two experiments, each using three types of sounds	1		precursor	0.333	other
	2	3	target A	0.333	other
	2		target B	0.333	other
	2		precursor	0.333	other

Note. Each experiment received a single point, which we distributed equally amongst the functional categories of the sounds used.

Classification of Temporal Structure

We classified sounds into one of five categories: (1) *flat*, (2) *percussive*, (3) *click train*, (4) *other*, and (5) *undefined* based on the descriptions given in the article and listed online links. We classified sounds as “flat” when their description included a period of invariant sustain and defined rise/fall times, such as “a 500-Hz sinusoid, 150msec in duration… gated with a rise-decay time of 25msec” (Watson & Clopton, 1969). Similarly, we classified sounds described as “rectangularly gated” (Robinson, 1973), having a “rectangular envelope” (Franěk, Mates, Radil, Beck, & Pöppel, 1991) or “abrupt onsets and offsets” (Hübner & Hafter, 1995) as flat. Samples of sounds falling into this category appear in Figure 2a (further information regarding flat tone offsets appear in Supplemental Figure 1).

Our second category of “percussive” encompassed sounds with sharp onsets followed by gradual decays with no sustain period (i.e. impact sounds). This included sounds from cowbells (Pfordresher & Palmer, 2006), bongos (Radeau & Bertelson, 1978), chimes and bells (Gregg & Samuel, 2009), and pianos (Pfordresher & Palmer, 2006) – in which hammers impact strings. Environmental impact sounds such as hand claps (Gregg & Samuel, 2009), footsteps (Pastore, Flint, Gaston, & Solomon, 2008) and dropped objects (Grassi, 2005) also fell within this category. Depictions of sound categorized as percussive are seen in Figure 2d and are summarized in detail in Supplemental Table 1b.

Our third category of “click train[1]” contained sounds described as of a series of repeated stimuli over a short duration. This included sounds explicitly identified as “click trains” (Shinn-Cunningham, 2000) “pulse trains” (Pollack, 1971) or “pulses in a train”

[1] Note: As they were less common in Music Perception, our teams’ earlier survey classified clicks and click trains within the “other” category.

[2] 3Hz amplitude modulated tones and amplitude modulated pedestal tones exhibit descriptions of temporal structures similar to flat tones with defined rise/fall times. In between rise and fall these tones contain some variation in amplitude. However, this

(Richardson & Frost, 1979), as well as those described as “a series of free-field acoustic clicks” (Uttal & Smith, 1967).

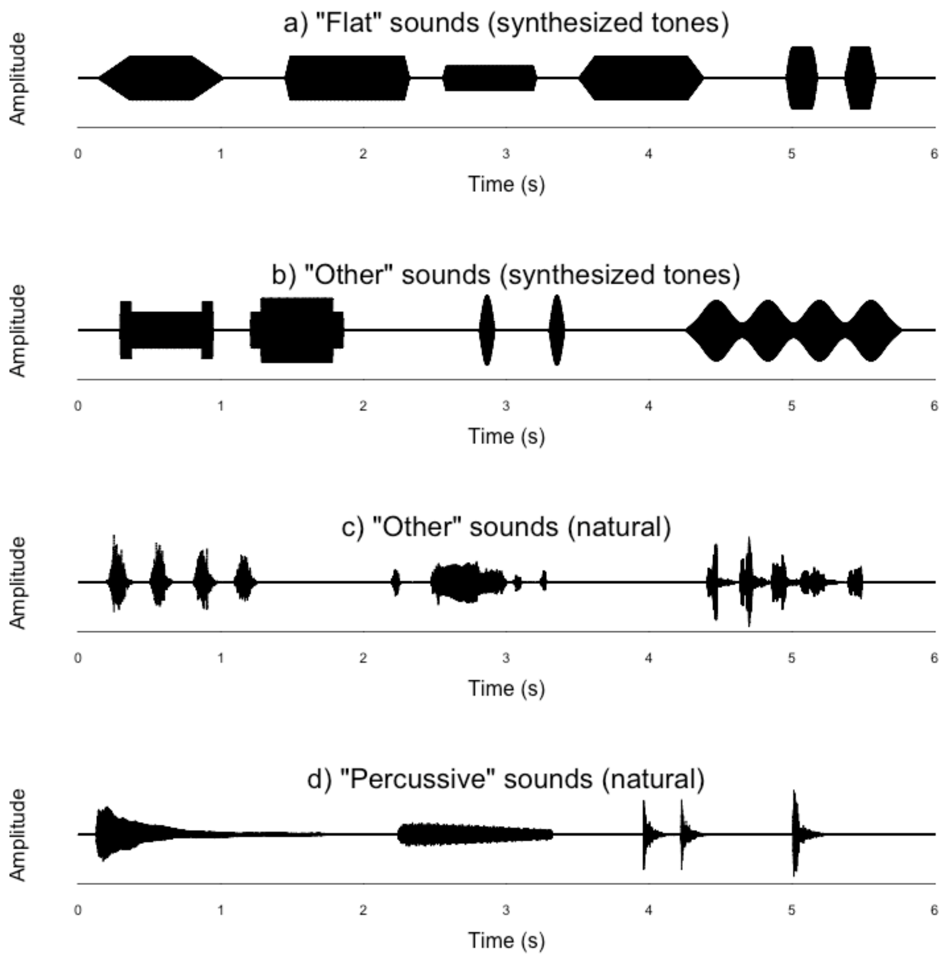


Figure 2. Examples of temporal structures encountered: (a) *flat* tones with assorted rise/fall times (the last two depict a 200ms tone shaped with 40ms cosine and linear ramps respectively), (b) synthesized tones in the *other* category including pedestal tones, “speedbump tones” and 3Hz amplitude modulated tones, (c) natural sounds in the *other* category such as a dog barking, chicken clucking and bird chirping, and (d) *percussive* sounds including a piano, bell, hand claps and bongo drum.

Our fourth category of “other” contained sounds with specified temporal structures other than those previously defined. This included 3Hz amplitude modulated tones (Riecke, van Opstal, & Formisano, 2008), amplitude modulated pedestal tones[2] (Bonnel & Hafter, 1998), tones with defined rise/fall times and no sustain (Hasuo, Nakajima, Osawa, & Fujishima, 2012; Wright & Fitzgerald, 2004), and complex environmental sounds (Gregg & Samuel, 2009) as well as brass (Gregg & Samuel, 2009; Stilp, Alexander, Kiefte, & Kluender, 2010), string (Stilp et al., 2010), and wind instruments (Gates, Bradshaw, & Nettleton, 1974). Supplemental Table 1a summarizes all sounds classified under this category and depictions of artificial and natural sounds are displayed in Figures 2b and 2c, respectively.

Finally, we used a fifth category of "undefined" for sounds whose temporal structures could not be discerned from the information provided (including online links, if given). For example, we classified sounds described as merely 'a 500ms 1000Hz tone' as undefined.

Additional Classifications

We also coded other characteristics such as spectral structure, duration, and intensity, as well as technical equipment information such as headphone/speaker model and tone generator for all stimuli. This builds on our team's previous approach (Schutz & Vaisberg, 2014) of classifying these properties only for stimuli with undefined temporal structures. In addition to contextualizing our findings by exploring differences in the degrees of definition of multiple stimulus properties, it provides further insight into research approaches in auditory perception.

[2] 3Hz amplitude modulated tones and amplitude modulated pedestal tones exhibit descriptions of temporal structures similar to flat tones with defined rise/fall times. In between rise and fall these tones contain some variation in amplitude. However, this variation is simplistic in comparison to the temporal structures of natural sounds.

Results & Discussion

Temporal Structure

As seen in Figure 3a[3] 53.9% of the sounds had undefined temporal structures, more than in the journal *Music Perception* (35%). This lack of definition was not confined to merely the shortest tones, but appears across all durations (see Supplemental Figure 2 for details). However, we suspect most sounds with undefined temporal structures had "flat" envelopes, in which case flat tones would comprise 84% of this survey, and 93% of stimuli surveyed lack the types of dynamic amplitude changes characteristic of natural sounds[4].

Given our sensitivity to sounds' temporal dynamics (Klatzky et al., 2000; Lutfi, 2007), we consider this documentation of the disproportionate focus on amplitude invariant tones to be the survey's most important outcome. Qualitative differences in performance based on the use of sounds with dynamically varying vs. amplitude invariant (Schutz, 2009; Vallet et al., in press) and even event-congruent vs. incongruent (i.e. Grassi & Casco, 2009) temporal structures demonstrate that conclusions drawn from experiments employing one temporal structure do not necessarily generalize to others. Although experiments employing "ecologically invalid" stimuli can be informative

[3] To ensure sounds such as background/masking noise did not artificially inflate the *undefined* category, we analyzed our results in three ways: (1) including background and masking noise, (2) analyzing them separately, and (3) removing them entirely. As these considerations did not meaningfully alter the outcome (Gillard & Schutz, 2013), here our analysis includes all sounds encountered.

[4] Including all 'undefined', 'flat', and 'click train' sounds, and select 'other' sounds – tones with rise/fall times and no sustain, amplitude-modulated pedestal tones and amplitude-modulated tones with modulators of 3Hz.
Percussive sounds account for 4.5% of the survey, and approximately 2.5% of 'other' sounds contain natural offsets such as musical instruments, environmental sounds, etc.

(Mook, 1983), we find this disconnect between the pervasive nature of temporally rich sounds outside the lab and the paucity of such sounds within important to consider. Even when seeking a theoretical understanding of the perceptual system's structure and function, it can be problematic to focus overwhelmingly on simplified stimuli. For example, it is well recognized that motion plays a crucial role in visual perception (Gibson, 1954), and experiments with static images provide only a partial understanding of visual processing – regardless of whether a study's goals were theoretical or applied.

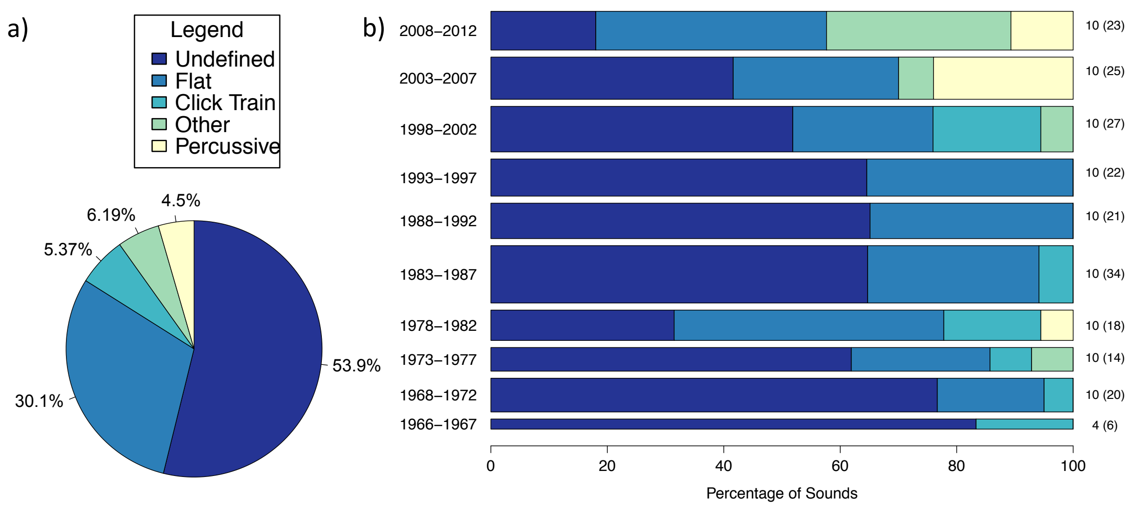


Figure 3. The temporal structures of surveyed sounds. Panel A gives a summary for all years. Panel B breaks down the results in 5-year periods, with bar height representing the number of points (i.e. experiments) contained. The number of articles in each period is displayed on the right vertical axis, with the number of points in brackets.

Temporal Structure Specification over Time

As seen in Figure 3b, specification of temporal structure has improved over time, rising from under 20% in 1966-1967 to just over 80% in 2008-2012. The use of tones with natural temporal shapes (i.e. percussive tones as well as instrumental and environmental sounds in the other category) has increased as well, averaging 26.7% of the sounds used between 2003 and 2012. Although this suggests increasing awareness of temporal structure's importance, sounds with simplistic temporal structures are still pervasive. Even within the most recent time window 57.6% of stimuli are presumed flat (i.e. flat + undefined) and only 29.6% exhibit the kinds of dynamically varying temporal structures characteristic of natural sounds. Furthermore, stimuli with dynamic structures were generally employed only in experiments explicitly assessing identification of natural events such as walking (Pastore et al., 2008), everyday sound recognition (Gregg & Samuel, 2009) impacts (Grassi, 2005), or musical tasks (Pfordresher & Palmer, 2006). Consequently amplitude invariant sounds clearly serve as the "default" stimulus within this corpus of stimuli, possibly biasing conclusions regarding the auditory system's structure and function.

General Specification of Stimulus Properties

The omission of a stimulus' temporal structure did not indicate general neglect of detail, as all other properties surveyed were specified at significantly higher rates (Table 2). Therefore, we believe this disproportionate omission of temporal structure suggests that it has not previously been thought to play an important role in experimental outcomes – or at least a role of lesser importance than the specific model of tone generator or headphone used to deliver a sound. Additional details regarding headphones/speakers and technical equipment can be found in Supplemental Table 2 and Supplemental Table 3 respectively.

Table 2

Percentage of defined properties

Property	Point Weighting	Percentage Defined	Binomial Sign Test
Duration	198.15	94.4%	$p < 0.001$
Spectral Structure	185.67	88.4%[5]	$p < 0.001$
Intensity	177.50	84.5%	$p < 0.001$
Headphones/Speakers	133.00	63.3%	$p < 0.001$
Sound Generator/Source	128.16	61.0%	$p < 0.001$
Temporal Structure	96.89	46.1%	-

Note. The binomial sign test was conducted assuming the rate of all defined properties is consistent with the rate of defined temporal structures (i.e. 46%).

Implications

Presuming tones with undefined temporal structures were in fact flat, it is clear that amplitude invariant temporal sounds are favored in auditory research. Although they offer a high degree of methodological control, it is evident that conclusions drawn with one type of temporal structure do not always generalize to others. This is evident in explorations of duration matching (Grassi & Darwin, 2006; Grassi & Pavan, 2012), duration discrimination (Schlauch et al., 2001), loudness matching (Ries et al., 2008) and loudness change discrimination (Neuhoff, 1998, 2001), as well as audio-visual integration (Schutz, 2009) and duration estimation (Vallet et al., in press). Therefore disproportionate focus on amplitude invariant stimuli can compel conclusions and

[5] 'Defined' spectral structures included explicit descriptions (i.e. pure tone, complex tone, white noise, etc.) as well as descriptions of 'tones' with a specified fundamental (i.e. "a 440 Hz tone"). These comprised 19.3% of sounds encountered, and presumably referred to pure tones.

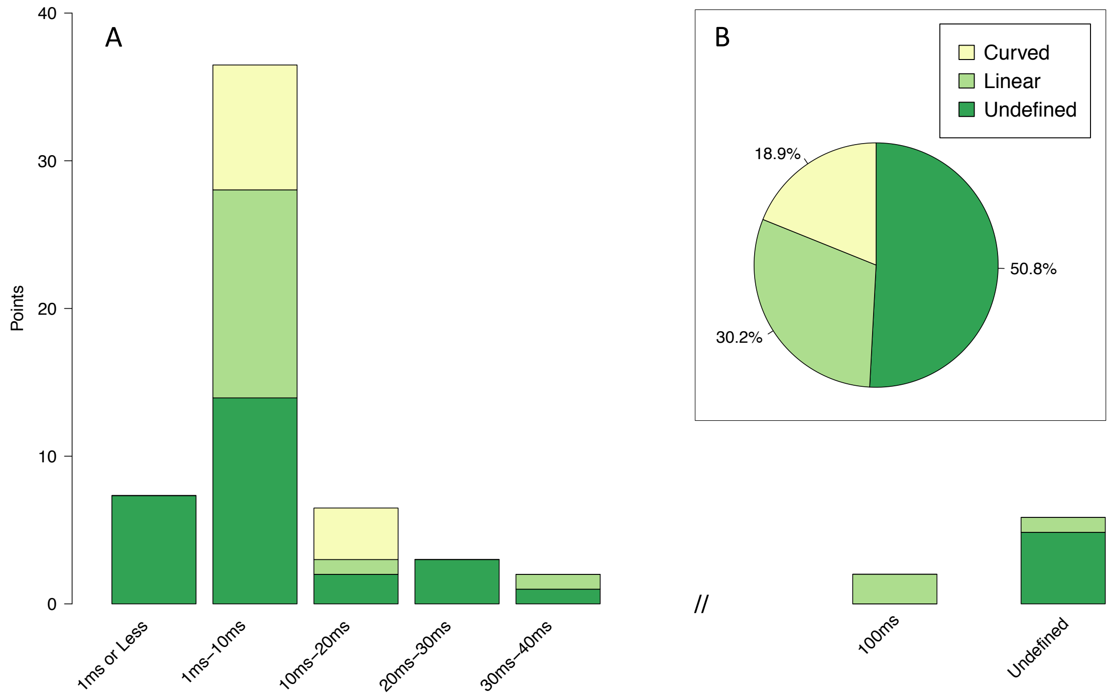
theories failing to generalize to sounds with natural, dynamic amplitude changes. For example, temporal structure plays a crucial role in triggering the 'identity decision' – perceiving two sensory inputs as originating from the same event (Bedford, 2001a, 2001b, 2004) – leading to decidedly different patterns of audio-visual integration when hearing sounds with natural/dynamic vs. artificial/amplitude invariant temporal structures (Schutz, 2009).

In conclusion, this empirical survey of sound characteristics builds upon previous speculative concerns voiced by prominent figures (i.e. Neuhoff, Plomp; as discussed by Neuhoff, (2004)), and long-standing concerns surrounding the disconnect between the sounds heard inside and outside the lab (Gaver, 1993a, 1993b). This survey extends previous findings by our research team examining the journal *Music Perception*, demonstrating this disproportionate focus on flat sounds is not limited to a single journal, domain of focus (i.e. music vs. general auditory perception), or paradigm. Consequently we hope this work will highlight the potential for future research on temporal structure, and we have posted a free software tool for generating sounds with dynamic temporal structures at www.maplelab.net/software. These findings are of relevance to both auditory cognition and psychoacoustics – two fields that historically have been quite divided (Neuhoff, 2004). Furthermore, they provide empirical support for previous observations and concerns that our understanding of the auditory system will only improve by using stimuli that "involve modulation [i.e. changes in temporal structure] in ways that are closer to real-world tasks faced by the auditory system." (Joris, Schreiner, & Rees, 2004).

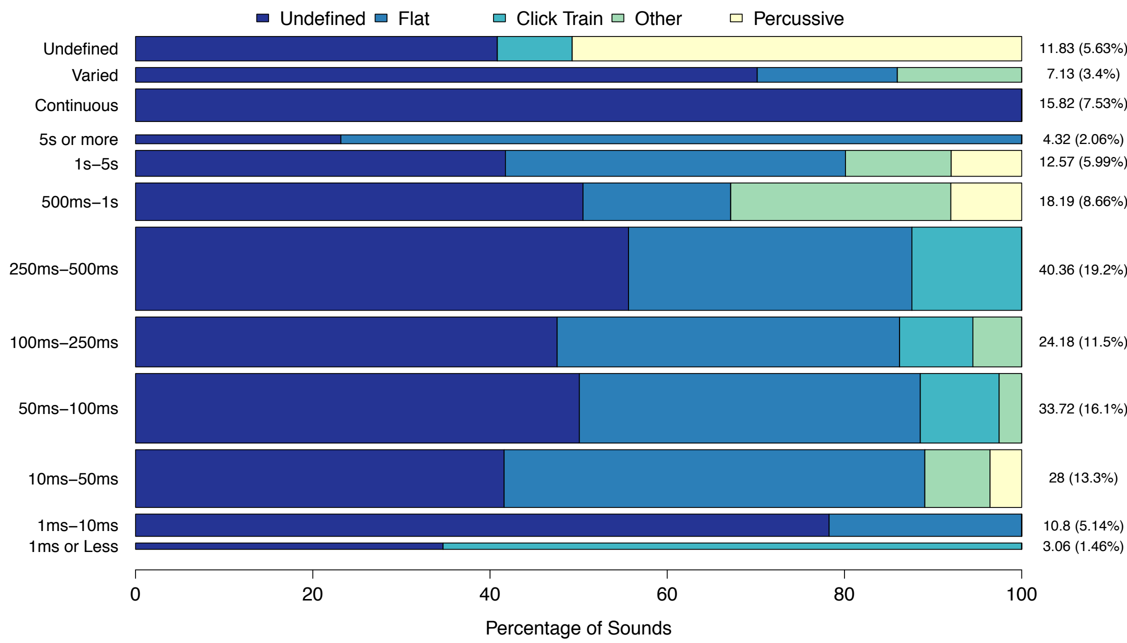
Acknowledgements

We would like to thank Martha Teghtsoonian, Peter Pfordresher and a third anonymous reviewer for their comments on an earlier version of this manuscript. Funding for this research came through grants from the Natural Sciences and Engineering Research Council of Canada (NSERC RGPIN/386603-2010), Ontario Early Researcher Award (ER10-07-195), Canadian Foundation for Innovation (CFI-LOF 30101), and the McMaster University Arts Research Board to Dr. Michael Schutz.

Supplemental Material



Supplemental Figure 1. Flat tone offset ramp types. Panel A summarizes prevalence of ramp type by offset duration and Panel B summarizes the distribution of ramp types within the flat category. Median offset duration was 9ms, with 76.5% of defined offset durations ≤ 10ms and 87.8% ≤ 20ms. Undefined ramp types were most likely linear, which would account for 81% of flat tone offsets. Curved offsets include cosine, cosine squared, raised cosine, quarter sine-wave and s-shaped ramps.



Supplemental Figure 2. Temporal structures by stimulus duration. Bar height represents the number of points (i.e. experiments) contained in each duration category. The number of points for each category is displayed on the right vertical axis, with the percentage of total points in brackets.

Supplemental Table 1

Detailed listing of sounds within the 'Other' and 'Percussive' categories

Sound	Point Weighting	% of Total Points
	(a) Other	
Pedestal Tones	1.5	0.71%
Speedbump Tones	2.918	1.39%
AM Tones (3Hz Modulator)	1.5	0.71%
Environmental sound – bird	0.724	0.34%
Environmental sound – dog	0.362	0.17%
Environmental sound – phone	0.181	0.09%
Environmental sound – ship	0.362	0.17%
Environmental sound – train	0.362	0.17%
Trumpet	0.362	0.17%
Saxophone	0.666	0.32%
French Horn	0.666	0.32%
Musical Excerpt	0.666	0.32%
Organ	1	0.48%
Total	**11.269**	**5.37%**
	(b) Percussive	
Bongos	1	0.48%
Cowbell	1.5	0.71%
Bell	0.4	0.19%
Chimes	0.362	0.17%
Piano	1.5	0.71%
Music Box	0.256	0.12%
Phone (bell) ring	0.181	0.09%
Hand Clap	0.256	0.12%
Footsteps	1	0.48%
Dropped Objects	3	1.43%
Total	**9.455**	**4.50%**

Supplemental Table 2

Detailed listing of headphones and speakers encountered

Brand	Point Weighting	% of Total Points
Headphones		
AKG Acoustics	8	3.81%
Belltone	1	0.48%
Beyer	7	3.33%
Grason-Stadler	8	3.81%
Koss	7	3.33%
Permoflux Corporation PDR	3	1.43%
Realistic Nova	2	0.95%
Sennheiser	14	6.67%
Sony	9	4.29%
Stax	1	0.48%
Telephonics Dynamic Headphone (TDH)	45	21.43%
Yamaha	9	4.29%
Custom	2	0.95%
Speakers		
Acoustic Research (AR)	2	0.95%
Audax	1	0.48%
Haram/Kardon	4	1.90%
Philips	1	0.48%
Radio Spares (RS)	3	1.43%
Revox	1	0.48%
Sony	1	0.48%
Undefined	77	36.67%
N/A	4	1.90%
Total	**210**	**100.00%**

Note. Defined headphones and speakers account for 55.42% and 6.19% of total points respectively. One study dropping balls on plates (accounting for 3 points) and another using custom-made buzzers (1 point) are represented in the N/A category, as they employed neither headphones nor speakers.

Supplemental Table 3

Detailed Listing of technical equipment used for sound creation

Method of Sound Generation	Point Weighting	% of Total Points
Technical Equipment - Software		
Cool Edit	4	1.90%
MATLAB	3	1.43%
IRCAM Musical Workstation	1	0.48%
MITSYN	7	3.33%
Praat Software	3	1.43%
SoundEdit	4	1.90%
Technical Equipment – Hardware (Computers)		
Commodore Computer	4	1.90%
Hewlett-Packard Personal Computer	3	1.43%
IBM Computer	10.5	5.00%
Phillips Minicomputer	1	0.48%
Olivetti Microcomputer	1	0.48%
Technical Equipment – Hardware (Tone Generators)[6]		
Advanced Audio-Signal Generator	2	0.95%
Belltone Audiometer	1	0.48%
Coulbourn Instruments Modules	6	2.86%
Elgenco Generator	0.666	0.32%
General Radio Generator	2.166	1.03%
Grason-Stadler Generator	13.25	6.31%
Grass Generator	3	1.43%
Hewlett-Packard Oscillator	3.833	1.83%
Klaat Synthesizer	2	0.95%
Krohn-Hite Oscillator	7.585	3.61%
Lafayette Instrument Co. Generator	1	0.48%
Madsen Audiometer	2	0.95%
MED Association Generator	1	0.48%
National Semiconductor Digital Noise Source	1	0.48%

[6] Specification rates for the temporal structures of stimuli created with such equipment (45.2%) was similar to overall specification levels (46.1%). Additionally, although some external tone generators may only be capable of producing flat tones, only 35% of stimuli with undefined temporal structures were created with such equipment.

Novatech Digital Synthesizer	1	0.48%
Schlumberger Generator	4	1.90%
Tucker Davis Processor/Converter	4	1.90%
Wavetek Oscillator	14.333	6.83%
Custom Generator	2	0.95%
Other Methods		
Electronic Organ	1	0.48%
Free-field Impact Event	3	1.43%
McGill University Musical Samples	1.333	0.63%
Recording – Bongos	1	0.48%
Recording – Environmental Sounds	3.808	1.81%
Recording – Footsteps	1	0.48%
Recoding – Musical Excerpt	0.666	0.32%
Roland Keyboard	3	1.43%
Undefined	81.86	38.98%
Total	**210**	**100%**

Note. Technical equipment included any specified hardware or software to generate the stimuli, such as tone generators, computers and computer software and accounted for 53.97% of total points. Other methods of sound generation, including recordings of sounds, sounds produced by musical instruments and sounds generated in free-field accounted for 7.05% of total points.

References

Armontrout, J. A., Schutz, M., & Kubovy, M. (2009). Visual determinants of a cross-modal illusion. *Attention, Perception, & Psychophysics*, *71*(7), 1618–1627.

Bedford, F. L. (2001a). Object identity theory and the nature of general laws: Commentary reply. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *20*(3-4), 277–293.

Bedford, F. L. (2001b). Towards a general law of numerical/object identity. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *20*(3-4), 113–175.

Bedford, F. L. (2004). Analysis of a constraint on perception, cognition, and development: one object, one place, one time. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 907–912.

Boltz, M., Marshburn, E., Jones, M. R., & Johnson, W. W. (1985). Serial-pattern structure and temporal-order recognition. *Perception & Psychophysics*, *37*(3), 209–217.

Bonnel, A., & Hafter, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Perception & Psychophysics*, *60*(2), 179–190.

DiGiovanni, J. J., & Schlauch, R. S. (2007). Mechanisms Responsible for Differences in Perceived Duration for Rising-Intensity and Falling-Intensity Sounds. *Ecological Psychology*, *19*(3), 239–264.

Franěk, M., Mates, J., Radil, T., Beck, K., & Pöppel, E. (1991). Sensorimotor synchronization: Motor responses to regular auditory patterns. *Perception & Psychophysics*, *49*(6), 509–516.

Gates, A., Bradshaw, J. L., & Nettleton, N. C. (1974). Effect of different delayed auditory feedback intervals on a music performance task. *Perception & Psychophysics*, *15*(1), 21–25.

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

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279–325.

Gibson, J. J. (1954). The visual perception of objective motion and subjective movement. *Psychological Review*, *61*(5), 304–314.

Gillard, J., & Schutz, M. (2013). The importance of amplitude envelope: Surveying the temporal structure of sounds in perceptual research. In *Proceedings of the Sound and Music Computing Conference* (pp. 62–68). Stockholm, Sweden.

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

Grassi, M. (2010). Sex difference in subjective duration of looming and receding sounds. *Perception*, *39*(10), 1424–1426.

Grassi, M., & Casco, C. (2009). Audiovisual bounce-inducing effect: attention alone does not explain why the discs are bouncing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 235–243.

Grassi, M., & Darwin, C. J. (2006). The subjective duration of ramped and damped sounds. *Perception & Psychophysics*, *68*(8), 1382–1392.

Grassi, M., & Pavan, A. (2012). The subjective duration of audiovisual looming and receding stimuli. *Attention, Perception & Psychophysics*, *74*(6), 1321–33.

Gregg, M., & Samuel, A. (2009). The importance of semantics in auditory representations. *Attention, Perception, & Psychophysics*, *71*(3), 607–619.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological Science*, *16*(3), 228–235.

Hasuo, E., Nakajima, Y., Osawa, S., & Fujishima, H. (2012). Effects of temporal shapes of sound markers on the perception of interonset time intervals. *Attention, Perception & Psychophysics*, *74*(2), 430–445.

Hübner, R., & Hafter, E. R. (1995). Cuing mechanisms in auditory signal detection. *Perception & Psychophysics*, *57*(2), 197–202.

Jackman, S. (2005). Pooling the polls over an election campaign. *Australian Journal of Political Science*, *40*(4), 499–517.

Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews*, *84*(2), 541–577.

Klatzky, R. L., Pai, D. K., & Krotkov, E. P. (2000). Perception of material from contact sounds. *Presence: Teleoperators and Virtual Environments*, *9*(4), 399–410.

Lutfi, R. A. (2007). Human Sound Source Identification. In W. A. Yost, R. R. Fay, & A. N. Popper (Eds.), *Auditory Perception of Sound Sources* (pp. 13–42). New York, NY: Springer.

Machado, A., & Keen, R. (1999). Learning to Time (LET) or Scalar Expectancy Theory (SET)? A Critical Test of Two Models of Timing. *Psychological Science*, *10*(3), 285–290.

Mook, D. (1983). In defense of external invalidity. *American Psychologist*, *38*(4), 379–387.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123–124.

Neuhoff, J. G. (2001). An Adaptive Bias in the Perception of Looming Auditory Motion. *Ecological Psychology*, *13*(2), 87–110.

Neuhoff, J. G. (2004). *Ecological Psychoacoustics*. (J. G. Neuhoff, Ed.) (pp. 1–13). Amsterdam: Elsevier Academic Press.

Pastore, R. E., Flint, J. D., Gaston, J. R., & Solomon, M. J. (2008). Auditory event perception : The source – perception loop for posture in human gait. *Perception & Psychophysics*, *70*(1), 13–29.

Pfordresher, P. Q., & Palmer, C. (2006). Effects of hearing the past, present, or future during music performance. *Perception & Psychophysics*, *68*(3), 362–376.

Pollack, I. (1971). Discrimination of restrictions in sequentially blocked auditory displays : Shifting block designs. *Perception & Psychophysics*, *9*(3B), 335–338.

Radeau, M., & Bertelson, P. (1978). Cognitive factors and adaptation to auditory-visual discordance. *Perception & Psychophysics*, *23*(4), 341–343.

Richardson, B. L., & Frost, B. T. (1979). Tactile localization of the direction and distance of sounds. *Perception & Psychophysics*, *25*(4), 336–344.

Riecke, L., van Opstal, A. J., & Formisano, E. (2008). The auditory continuity illusion: A parametric investigation and filter model. *Perception & Psychophysics*, *70*(1), 1–12.

Ries, D. T., Schlauch, R. S., & DiGiovanni, J. J. (2008). The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds. *The Journal of the Acoustical Society of America*, *124*(6), 3772–3783.

Robinson, C. E. (1973). Reaction time to the offset of brief auditory stimuli. *Perception & Psychophysics*, *13*(2), 281–283.

Scheuren, B. F. (2004). *What is a survey? American Statistical Association*. Retrieved from http://www.whatisasurvey.info/

Schlauch, R. S., Ries, D. T., & DiGiovanni, J. J. (2001). Duration discrimination and subjective duration for ramped and damped sounds. *The Journal of the Acoustical Society of America*, *109*(6), 2880–2887.

Schutz, M. (2009). *Crossmodal integration: The search for unity*. University of Virginia.

Schutz, M., & Kubovy, M. (2009a). Causality and cross-modal integration. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1791–1810.

Schutz, M., & Kubovy, M. (2009b). Deconstructing a musical illusion: Point-light representations capture salient properties of impact motions. *Canadian Acoustics*, *37*(1), 23–28.

Schutz, M., & Lipscomb, S. (2007). Hearing gestures, seeing music: Vision influences perceived tone duration. *Perception*, *36*(6), 888–897.

Schutz, M., Stefanucci, J. K., Carberry, A., & Roth, A. (under review). Name that percussive tune: Associative memory and amplitude envelope.

Schutz, M., & Vaisberg, J. (2014). Surveying the temporal structure of sounds used in Music Perception. *Music Perception: An Interdisciplinary Journal*, *31*(3), 288–296.

Sekuler, R., Sekuler, A. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308. Retrieved from http://www.mendeley.com/research/sound-alters-visual-motion-perception/

Shinn-Cunningham, B. (2000). Adapting to remapped auditory localization cues: a decision-theory model. *Perception & Psychophysics*, *61*(2), 33–47.

Smith, T. (1983). On the validity of inferences from non-random sample. *Journal of the Royal Statistical Society. Series A (General)*, *146*(4), 394–403.

Snow, J. C., Pettypiece, C. E., McAdam, T. D., McLean, A. D., Stroman, P. W., Goodale, M. A., & Culham, J. C. (2011). Bringing the real world into the fMRI scanner: repetition effects for pictures versus real objects. *Scientific Reports*, *1*(130), 1–10.

Stecker, G. C., & Hafter, E. R. (2000). An effect of temporal asymmetry on loudness. *The Journal of the Acoustical Society of America*, *107*(6), 3358–3368.

Stilp, C. E., Alexander, J. M., Kiefte, M., & Kluender, K. R. (2010). Auditory color constancy: Calibration to reliable spectral properties across nonspeech context and targets. *Attention, Perception, & Psychophysics*, *72*(2), 470–480.

Teghtsoonian, R., Teghtsoonian, M., & Canévet, G. (2005). Sweep-induced acceleration in loudness change and the “bias for rising intensities”. *Perception & Psychophysics*, *67*(4), 699–712.

Uttal, W. R., & Smith, P. (1967). Contralateral and heteromodal interaction effects in somatosensation: Do they exist? *Perception & Psychophysics*, *2*(8), 363–368.

Vallet, G., Shore, D., & Schutz, M. (in press). Exploring the role of amplitude envelope in duration estimation. *Perception*.

Walker, J., & Scott, K. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(6), 1327–1339.

Warren, 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.

Watson, C. S., & Clopton, B. M. (1969). Motivated changes of auditory sensitivity in a simple detection task. *Perception & Psychophysics*, *5*(5), 281–287.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667.

Wright, B. A., & Fitzgerald, M. B. (2004). The time course of attention in a simple auditory detection task. *Perception & Psychophysics*, *66*(3), 508–516.

Appendix

SURVEY FUNCTIONS DOCUMENTATION

User Functions

prettySurveyFunction

Description
Removes all pilot data, speech studies, unclassified envelopes and empty rows from the data and changes the envelope category 'Not Specified' to 'Unspecified'

Usage
prettySurveyFunction (dat)

Arguments
dat - The initial data frame containing all survey information

Examples
APP2=prettySurveyFunction(APP)

rearrangeAggregate

Description
Reorders the envelope or spectral categories in aggregate data frame by calling another function (i.e. rearrangeEnvelope or rearrangeSpectral), which contains specific order information.

Usage
rearrangeAggregate (Agg)

Arguments
Agg – the aggregate data frame containing envelope or spectral information

Examples
```
# Create aggregate data frame (default is to organize in alphabetical order)
appAgg = aggregate(APP2$Point.Weighting, list(APP2$Envelope), FUN=sum)
colnames (appAgg)=c('envelope', 'weighting')
# Rearrange aggregate to custom order
appAgg=rearrangeAggregate(appAgg)
```

Original appAgg			Rearranged appAgg		
	envelope	weighting		envelope	weighting
1	Click	2.00000	1	Unspecified	113.09033
2	Click Train	11.00000	2	Flat	63.16767
3	Flat	63.16767	3	Click	2.00000
4	Other	11.26900	4	Click Train	11.00000
5	Percussive	9.45800	5	Other	11.26900
6	Unspecified	113.09033	6	Percussive	9.45800

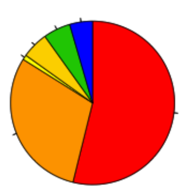
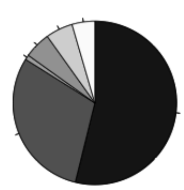
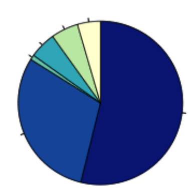
getColours

Description
Retrieves a list of colours for plotting.

Usage
getColours (Agg, col='orig')

Arguments
Agg – the aggregate data frame containing envelope or spectral information
col – a string representing the colour scheme to be used. Default is ‘orig’, but other options include ‘bw’, ‘pastel’ and ‘pf’.
‘orig’ – original color scheme using red, orange, yellow, green and blue
‘bw’ – black & white color scheme, using various hues of grey
‘pastel’ – pastel color scheme using Purple, Cyan, Yellow, Green and Pink
‘pf’ – printer friendly color scheme using Jacksons Purple, Jelly Bean, Pelorous, Moss Green and Cream. Note: This color scheme is coded in HEX.

Examples
```
pie(appAgg$weighting,col=getColours(appAgg)) # will use original colour scheme
pie(appAgg$weighting,col=getColours(appAgg, col='bw')) # black & white version
pie(appAgg$weighting,col=getColours(appAgg, col='pastel')) # pastel version
pie(appAgg$weighting,col=getColours(appAgg, col='pf')) # printer friendly version
```

‘orig’	‘bw’	‘pastel’	‘pf’
			

getDensity

Description
Retrieves a list of densities for plotting. This creates sections with hatching.

Usage
getDensity (Agg)

Arguments
Agg – the aggregate data frame containing envelope or spectral information

Examples
pie(appAgg$weighting, density=getDensity (appAgg))

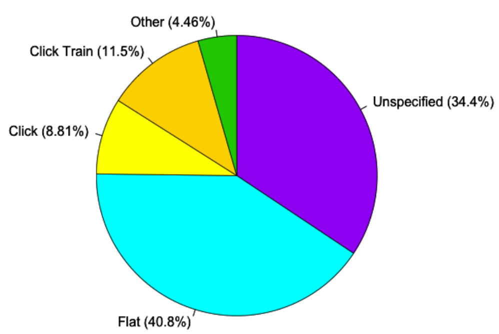
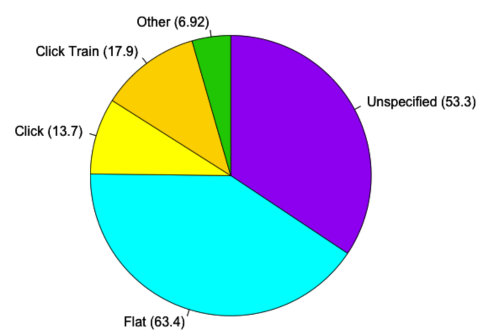
getLabel

Description
Returns a list of labels for plotting.

Usage
getLabel (Agg, attribute='envelope', type='percent')

Arguments
Agg – the aggregate data frame containing envelope or spectral information
attribute – 'envelope' or 'spectral' (i.e. name of first column in aggregate data frame). The default is 'envelope'
Type – 'percent' or 'absolute' to display percentages or absolute point values respectively. The default is 'percent'

Examples
# Pie chart displaying percentages
pie(hrAgg$weighting, col=getColours(hrAgg, col='pastel'),getLabel(hrAgg))
# Pie chart displaying points
pie(hrAgg$weighting, col=getColours(hrAgg, col='pastel'),getLabel(hrAgg, type='absolute'))

<table>
<tr><th>Percentage</th><th>Points</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

sumByParameter

Description
Returns a matrix containing sums of an attribute (i.e. envelope or spectral categories) based on set parameter (i.e. year window, range of durations, etc.).

Usage
sumByParameter (dat, bins, attribute="Envelope", parameter="yearGroup")

Arguments
dat – the data frame containing columns with raw points and the attribute of interest
bins – the number of bins required to make the data frame (i.e. the total # of years/range of each bin or max index number for a duration plot)
attribute – the name of the column contain the variable of interest (i.e "Envelope", "Simple.Spectral"). The default is ‘Envelope’.
parameter – the name of the newly created column that will contain the variable groupings (i.e. yearGroup, durIndex, etc.)

Examples
app=sumByParameter(APP2, bins=10)
dur1=sumByParameter(sDur, bins=9, parameter='durIndex')

yearGrouping

Description
Adds the column ‘yrBin’ to the data frame, which contains the range of years for each index bin number (i.e. 1998-2002, 2003-2007, etc.)

Usage
yearGrouping (dat, yearBin, startingYear='default')

Arguments
dat – the data frame containing columns with raw points and years
yearBin – the range of years covered by each bar in the plot (i.e. 5 year increments)
startingYear – The first year going backwards in time to include in plot (i.e. 2012, 2001, etc.). The default is max(Year), but can be specified (i.e. 2009)

Examples
APP2=yearGrouping(APP2, yearBin=5)
MP2=yearGrouping(MP, yearBin=5, startingYear=2008)

refineNonTargetNoise

Description
Refines the types of non-target noise (i.e. Flat Noise or Unspecified Noise).

Usage
refineNonTargetNoise (dat, maskers=T, backgroundNoise=T)

Arguments
dat – the data frame containing columns with raw points, spectral and envelope information
maskers – Default is TRUE. If you DO NOT want to refine masker types, set this variable to FALSE.
backgroundNoise – Default is TRUE. If you DO NOT want to refine background noise types, set this variable to FALSE.

Examples
APP2= refineNonTargetNoise (APP) # refines both maskers & background noise
APP2= refineNonTargetNoise (APP, maskers=F) # refines only background noise
APP2= refineNonTargetNoise (APP, backgroundNoise =F) # refines only maskers

removeNonTargetNoise

Description
Removes the types of non-target noise (i.e. Flat Noise or Unspecified Noise).

Usage
refineNonTargetNoise (dat, maskers=T, backgroundNoise=T)

Arguments
dat – the data frame containing columns with raw points, spectral and envelope information
maskers – Default is TRUE. If you DO NOT want to refine masker types, set this variable to FALSE.
backgroundNoise – Default is TRUE. If you DO NOT want to refine background noise types, set this variable to FALSE.

Examples
APP2= refineNonTargetNoise (APP) # removes both maskers & background noise
APP2= refineNonTargetNoise (APP, maskers=F) # removes only background noise
APP2= refineNonTargetNoise (APP, backgroundNoise =F) # removes only maskers

defineBreakPoints

Description
Creates a column in the data frame to store index numbers at specified numeric break points.

Usage
defineBreakPoints (dat, columnName, parameter, listOfBreaks)

Arguments
dat – the data frame containing the parameter for which to create break points
columnName – the name of the column in which to store the break point index
parameter – the name of the column containing the parameter of interest
listOfBreaks – a list containing the numeric break points to use

Examples
sDur=defineBreakPoints(sDur, columnName='durIndex', parameter='dur', listOfBreaks=c(1,10,50,100,250,500,1000,5000))

prettyDefinedDuration

Description
Returns a subset of the data containing only specified (i.e.numeric) durations.

Usage
prettyDefinedDuration (dat)

Arguments
dat – The data frame that contains Simple Duration information.

Examples
sDur=prettyDefinedDuration(APP2)

prettyUndefinedDuration

Description
Returns a subset of the data containing only stimuli with Unspecified, Continuous and Variable durations.

Usage
prettyUndefinedDuration (dat)

Arguments
dat – The data frame that contains Simple Duration information.

Examples
nsDur=prettyUndefinedDuration(APP2)

articleSelectionCount

Description
Counts the number of articles extracted per year and places these counts in a data frame.

Usage
articleSelectionCount (dat, startingYr=min(dat$Year), endingYr=max(dat$Year))

Arguments
dat – the data frame containing year and study information
startingYr – the starting year. The default is the minimum year in the data, but can be set to any number.
endingYr – the ending year. The default is the maximum year in the data, but can be set to any number.

Examples
hrSelection=articleSelectionCount(HR2,1980,2012)
appSelection=articleSelectionCount(APP2,1965,2012)
mpSelection=articleSelectionCount(MP2,1983,2012)
hppSelection=articleSelectionCount(HPP2,1975,2012)

articleSelectionCountLego

Description
Transforms a single row count matrix of article selections per year into a Lego matrix.

Usage
articleSelectionCountLego (matrix)

Arguments
matrix – the matrix returned by the articleSelectionCount function

Examples
hrSelection=articleSelectionCount(HR2,1980,2012)
hrLego = articleSelectionCountLego(hrSelection)

appSelection=articleSelectionCount(APP2,1965,2012)
appLego = articleSelectionCountLego(appSelection)

getSpecifiedVs.Unspecified

Description
Returns the points and percentages of specified vs. unspecified parameters in an aggregate data frame.

Usage
getSpecifiedVs.Unspecified (dat, parameter='Simple.Spectral')

Arguments
dat – The data frame that contains raw points and information regarding the

parameter of interest
parameter – the column name of the parameter of interest (i.e. Envelope, Simple.Spectral, Simple.Duration, Intensity, Headphones.Speakers, Sound.Generator, etc.) These arguments must be entered as a string with quotation marks. The default is set to 'Simple.Spectral'.

Examples

specAgg=getSpecifiedVs.Unspecified(APP2)
hpAgg=getSpecifiedVs.Unspecified(APP2, parameter= 'Headphones.Speakers')
durAgg=getSpecifiedVs.Unspecified(APP2, parameter= 'Simple.Duration')

getRefinedSpecifiedVs.Unspecified

Description

Returns the points and percentages of temporally specified and unspecified sounds that specify and do not specify other parameters in an aggregate data fame.

Usage

getRefinedSpecifiedVs.Unspecified (dat, parameter='Simple.Spectral', numOfCategories=3)

Arguments

dat – The data frame that contains information regarding Envelope and the parameter of interest.
parameter - The parameter of interest. The default is set to 'Simple.Spectral', but can be changed to any other parameter of interest (i.e. 'Headphones.Speakers', 'Simple.Duration', 'Intensity', 'Sound.Generator', etc.). Note: these parametes must be entered as a string, using quotation marks as demonstrated above.
numOfCategories – the number of categories returned (i.e. 2, 3 or 4). The default is 3. Categories are organized by the specification of envelope followed by the parameter (i.e. Unspecified envelope/Unspecified parameter, Unspecified envelope/Specified parameter, etc.). Outputs are listed below:

4 - Unspecified/Unspecified, Unspecified/Specified, Specified/Unspecified, Specified/Specified
3 - Unspecified/Unspecified, Unspecified/Specified, Specified Temporal structure Only
2 - Unspecified/Unspecified, Unspecified/Specified, Specified

Examples

hpAgg=getRefinedSpecifiedVs.Unspecified(APP2, numOfCategories=2, parameter= 'Headphones.Speakers')
hpAgg=getRefinedSpecifiedVs.Unspecified(APP2, numOfCategories=3, parameter= 'Headphones.Speakers')
hpAgg=getRefinedSpecifiedVs.Unspecified(APP2, numOfCategories=4, parameter= 'Headphones.Speakers')

numOfCategories	Output			
	envelope/param	weighting	percent	percentOf
numOfCategories=2	Unspecified/Unspecified	46.44200	41.1	41.1%
	Unspecified/Specified	66.64833	58.9	58.9%
	envelope/param	weighting	percent	percentOf
numOfCategories=3	Unspecified/Unspecified	46.44200	22.1	41.1%
	Unspecified/Specified	66.64833	31.7	58.9%
	.	96.89467	46.1	.
	envelope/param	weighting	percent	percentOf
numOfCategories=4	Unspecified/Unspecified	46.44200	22.1	41.1%
	Unspecified/Specified	66.64833	31.7	58.9%
	Specified/Unspecified	30.49400	14.5	31.5%
	Specified/Specified	66.40067	31.6	68.5%

CHAPTER 4

GENERAL DISCUSSION

As discussed in Chapter 3, amplitude envelope has the ability to affect performance on a variety of tasks, including associative memory. Although we had observed percussive melody associations being learned faster and more associations being recalled overall in a previous study conducted by our research group (Schutz, Stefanucci, Carberry, & Roth, under review), we were not able to replicate this finding using medical alarm referents. It is possible that in order for certain sounds to improve or alter performance, there must be some inherent meaning or information conveyed by the sound. This has been observed in tasking involving audio-visual integration, loudness change and perceived duration.

In audio-visual integration, a causal link appears to be the deciding factor of whether our perceptual system does or does not integrate the audio and visual information (Schutz, 2009). This is supported by the fact that visual gestures influence duration judgments of percussive but not flat tones (Schutz & Kubovy, 2009; Schutz, 2009) and two disks crossing paths appear to bounce when paired with damped tones but not ramped tones (Grassi & Casco, 2009). Essentially, decaying sounds are caused when two objects collide, which promote integration with visual impact events.

Object motion can be signaled artificially by sounds increasing or decreasing in amplitude over time. Those increasing in amplitude, known as “looming” or “ramped” sounds, appear to approach the listener and are perceived to change more in loudness than equivalent “receding” or “damped” sounds that decrease in amplitude over time (Neuhoff, 1998, 2001). Furthermore, these approaching sounds are perceived to stop closer to the listener than receding sounds (Neuhoff, 2001). These findings have been attributed to an adaptation, in which an overestimation of approaching sound sources allows the listener to prepare for source contact (Neuhoff, 2001).

An asymmetry has also been reported in the perceived duration of ramped and damped sounds, with damped sounds consistently sounding shorter (Grassi & Pavan, 2012; Schlauch, Ries, & DiGiovanni, 2001). Some have suggested that the perceptual system may discount part of the damped sound as an echo (Stecker & Hafter, 2000). This theory has been partially supported by empirical evidence, as instructions to evaluate all aspects of the sound decreased the asymmetry, but did not eliminate it entirely (DiGiovanni & Schlauch, 2007).

Among these other tasks, the temporal structure appears to be conveying some sort of additional information that is lacking in tones with flat temporal structures. More specifically, a sound's decay can inform listeners of the materials creating the sound (Klatzky, Pai, & Krotkov, 2000; Lutfi, 2007), or the outcome of an impact, such as bouncing or breaking (Warren & Verbrugge, 1984).

Why then would we see an effect of amplitude envelope on associations of objects but no effect when using medical alarm referents? There are many possible explanations that require further investigation. One potential explanation we explored is that causal links may be stronger for concrete objects than abstract medical alarm referents (Gillard & Schutz, 2013). For example, some of the objects used such as a clock or camera naturally make percussive sounds, which might make some association easier to learn and recall. Theses types of connections are harder to make with referents such as 'Oxygen', 'Temperature' or 'Perfusion'. While we did find some evidence for this theory

(Gillard & Schutz, 2013), further pilot work (not included in the current thesis) did not fully support this hypothesis.

Regardless, we believe that amplitude envelope is a topic with great potential to increase our understanding of the auditory system. Furthermore, music cognition is an ideal sub-filed for this undertaking as it can merge information from ecological acoustics and the more traditional psychoacoustic approaches of auditory perception research. Such investigations can promote cross-talk between these two very different perspectives, fuel the circulation of new findings and spark new ideas.

References for Chapter 4

DiGiovanni, J. J., & Schlauch, R. S. (2007). Mechanisms Responsible for Differences in Perceived Duration for Rising-Intensity and Falling-Intensity Sounds. *Ecological Psychology*, *19*(3), 239–264.

Gillard, J., & Schutz, M. (2013). Bridging the gap between sound and meaning. In *Proceedings of the Society for Music Perception and Cognition Conference* (p. 28). Toronto, ON.

Grassi, M., & Casco, C. (2009). Audiovisual bounce-inducing effect: attention alone does not explain why the discs are bouncing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(1), 235–243.

Grassi, M., & Pavan, A. (2012). The subjective duration of audiovisual looming and receding stimuli. *Attention, Perception & Psychophysics*, *74*(6), 1321–1333.

Klatzky, R. L., Pai, D. K., & Krotkov, E. P. (2000). Perception of material from contact sounds. *Presence: Teleoperators and Virtual Environments*, *9*(4), 399–410.

Lutfi, R. A. (2007). Human Sound Source Identification. In W. A. Yost, R. R. Fay, & A. N. Popper (Eds.), *Auditory Perception of Sound Sources* (pp. 13–42). New York, NY: Springer.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123–124.

Neuhoff, J. G. (2001). An Adaptive Bias in the Perception of Looming Auditory Motion. *Ecological Psychology*, *13*(2), 87–110. doi:10.1207/S15326969ECO1302_2

Schlauch, R. S., Ries, D. T., & DiGiovanni, J. J. (2001). Duration discrimination and subjective duration for ramped and damped sounds. *The Journal of the Acoustical Society of America*, *109*(6), 2880–2887.

Schutz, M. (2009). *Crossmodal integration: The search for unity*. University of Virginia.

Schutz, M., & Kubovy, M. (2009). Causality and cross-modal integration. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1791–1810.

Schutz, M., Stefanucci, J. K., Carberry, A., & Roth, A. (under review). Name that percussive tune: Associative memory and amplitude envelope.

Stecker, G. C., & Hafter, E. R. (2000). An effect of temporal asymmetry on loudness. *The Journal of the Acoustical Society of America*, *107*(6), 3358–3368.

Warren, 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.