EXPLORING THE ROLE OF TEMPORAL VARIATION IN THE DETECTION AND SUBJECTIVE ANNOYANCE OF AUDITORY ALARM SIGNALS

By Liam Foley, B.Sc. (Hons)

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

McMaster University

Copyright by Liam Foley, September 2021

McMaster University MASTER OF SCIENCE (2021) Hamilton, Ontario (Psychology, Neuroscience, and Behaviour)

TITLE: Exploring the Role of Temporal Variation in the Detection and Subjective Annoyance of Auditory Alarm Signals

AUTHOR: Liam Foley, B.Sc. (Hons) (Memorial University of Newfoundland and Labrador)

SUPERVISOR: Dr. Michael Schutz

NUMBER OF PAGES: 67

Preface

This thesis is composed of two independent manuscripts for publication. The first (Chapter 2) "More Detectable, Less Annoying. The Role of Temporal Variation in Envelope and Spectral Content on Detection and Annoyance" will be submitted to Psychological Science. This manuscript explores how detection and annoyance of sound are affected by temporal variation in two acoustic parameters; amplitude envelope and spectral content. The second (Chapter 3) "Improving Detectability of Auditory Interfaces Through Temporal Variation in Envelope" will be submitted to Human Factors. Here I build off of the previous manuscript by investigating tone detection in a split attention task more pertinent to the normative use of auditory interfaces. The author of this thesis is the primary author of both papers.

Acknowledgements

I would first like to thank my supervisor Dr. Michael Schutz for the support and guidance over the last two years. This time has certainly brought many (many many many) challenges for doing any sort of research. I am thankful to have had a great supervisor to help guide my development as a researcher. I would like to thank my committee Dr. Patrick Bennett and Dr. Ian Bruce, your guidance during this project has been invaluable. Thanks to Dr. Joe Schlesinger. Your first-hand experience has given me important perspectives to real world alarm problems that has helped develop this project.

Next I would like to thank my good friend, lab mate, collaborator, and partner in science, Cameron Anderson. A pandemic is a weird time to be completing grad school, without your support and encouragement I do not think I would be where I am. From offering a quick proof read, a virtual drink, or a fresh perspective you have been an amazing friend. Thanks to Konrad Swierczek, for answering my numerous digital audio questions over the last couple years.

I would also like to thank the many MAPLE Lab members, past and present, who I have had the privilege of working with. This lab has had some of the brightest and most passionate students I have ever seen. Their resiliency in the face of the pandemic has been nothing short of inspiring. Without the help of QQ's and volunteers Sara Marshall, Kailey McMillan, Thomas Samson-William, Jessica LaMantia, Subeetsha Uthayakumar, Jen Li, William Zhang, JJ Magee, Vivian Li, Jeffrey Ong, and Andy Elizondo Lopez, this project would have gone nowhere! I have had the privilege of supervising two great honours students, Laura Schachtler and Subeetsha Uthayakumar, both of which helped further the auditory alarms project. I would also like to thank Ben Kelly, for being an amazing Lab Manger and friend. Thanks you for catching those double spaces. Julia Bissessar thank you for your help with 2MP3 and for being a great lab manager. Max Delle Grazie thanks for your endless knowledge and passion for music. I know you are going to do amazing things in grad school. A big thank you to Andy Elizondo Lopez for his help during my final semester. You have been a pleasure to work with and your passion for the project is going to bring you far.

To Sejla, Josh, and Thomas, you are incredible friends and have been such a help over the course of this degree. You have been supportive and encouraging and I thank you. Thank you to my parents who I owe so much to. From setting up life in Hamilton, to care packages, to helping us resettle in St. John's I am grateful to have had such supportive parents during this time.

Finally, to my loving partner Rhiannon. Thank you for your support, kindness, and advice over the rocky course of this degree. From navigating a new city, to moving in the middle of the pandemic, to teaching me various things in R, you have been my rock.

Table of Contents

List of Figures

Chapter 1

General Introduction

Sound plays an integral role in our understanding of the world. Today many of the sounds we interact with are designed, which can be manipulated to ensure a particular function in line with our auditory system. This thesis is composed of two projects exploring how temporal variation in acoustic parameters can improve these sonic interactions. These investigations uncover important findings on designing more effective and ergonomic sounds for areas such as auditory alarms. In addition, these investigations further our understanding on the role temporal variation plays in our auditory perception.

Chapter 2 explores the role of temporal variation in a sound's amplitude envelope and spectral content in both detection and annoyance of sounds. Participants detected temporally variant tones significantly more than invariant tones, and found them significantly less annoying. This study adds to a growing body of literature showing the importance of temporal variation in auditory perception.

Chapter 3 builds on this finding, exploring detection of temporally variable tones during a concurrent speech task. Participants again detected temporally variant tones significantly more often than invariant tones, and found them significantly less annoying. This study adds to growing evidence that temporally variable sounds can help improve efficacy of auditory alarms. Combined chapters 2 and 3 document the original research I have completed for my M.Sc.

Chapter 2

More Detectable, Less Annoying. The Role of Temporal Variation in Envelope and Spectral Content on Detection and Annoyance.

Liam Foley, Joseph Schlesinger, Michael Schutz

Abstract

Auditory interfaces, such as auditory alarms, alert us to important information in our lives. Unfortunately poor detectability and annoyance inhibit the efficacy of many alarms. In this experiment we explore the role of two acoustic properties, amplitude envelope, changes in a sounds energy over time, and spectral content, in both detection and annoyance of sound. Undergraduate students rated percussive tones as significantly less annoying than flat tones. Crucially, this annoyance reduction did not come with a detection cost, as participants detected percussive tones significantly better than flat tones (particularly at low listening levels). Additionally, we found reductions in the duration of spectral content significantly lowered annoyance ratings, without a commensurate reduction in detection. Together, these findings help inform our theoretical understanding of detection and annoyance of sound. In addition, these acoustic properties serve as promising new design considerations for auditory alarms.

Introduction

Our hearing allows for better understanding of our environment. Exploring what acoustic factors aide in parsing competing acoustic signals is integral to our understanding of the auditory system, as well as the design of effective auditory interfaces. One type of auditory interfaces whose problems have been extensively researched is the auditory alarm (Edworthy, 2013). They are commonly designed to sound often and loudly in an attempt to ensure detection (Edworthy & Hellier, 2005; Schlesinger et al., 2018); however, their numerosity leads to issues of annoyance and detection, reducing their usefulness (Rayo, Patterson, Abdel-Rasoul, & Moffatt-Bruce, 2019).

Recent investigations into the use of temporal variation in envelope, a relatively understudied acoustic property in perception (Schutz & Gillard, 2020), has shown promise in improving auditory alarm efficacy. In particular, it can be used to reduce annoyance without harming learning, or memory of alarms (Sreetharan, Schlesinger, & Schutz, 2021). Here, we build on that work by exploring how temporal variation in both amplitude envelope (flat and percussive envelopes) and spectral content (differing durations of harmonic energy) affect not only annoyance, but also a core issue of auditory perception—tone detection.

Auditory Alarms: Detectable Yet Annoying

Auditory alarms are invaluable in many safety critical industries, including (but not limited to) industrial processing (Laberge, Bullemer, Tolsma, & Reising, 2014), aviation (Bliss, 2003), automobiles, (Marshall, Lee, & Austria, 2007), railways (Edworthy, Hellier, Titchener, Naweed, & Roels, 2011) and healthcare (Edworthy, 2013). These alarms are typically easily detectable in isolation; however, common design philosophy dictates that loud alarms must sound often in order to be detected (Schlesinger et al., 2018) leading to an unmanageable sonic cacophony in their real world use (Sendelbach & Funk, 2013).

The sheer numerosity of alarms prevalent in both healthcare (Varpio, Kuziemsky, MacDonald, & King, 2012) and industrial settings (Laberge et al., 2014) exacerbates problems with loud volume levels. Studies of healthcare alarms have shown alarm rates of more than three hundred and fifty alarms per patient per day (Sendelbach & Funk, 2013). Other reports have shown only 0.5% indicate life threatening events (O'Carroll, 1986). Numerous alarms combined with low positive predictive value creates an unfortunate vicious cycle. Making alarms annoying helps reduce misses, but large numbers of attention-grabbing alarms incentivize users to tune them out or even turn them off— resulting in disastrous consequences (Bliss, Gilson, & Deaton, 1995; Block, Nuutinen, & Ballast, 1999), such as one patient who died despite a warning alarm sounding for over an hour (Cvach, 2012; Sendelbach & Funk, 2013).

Role of Temporal Structure in Recognition and Perception

Temporal variation in envelope is a prominent feature in natural sounds (Foley & Schutz, 2021). Musical sounds, for example, exhibit great temporal variation in overall envelope and in their constituent component frequencies (Schutz, 2019). These temporal fluctuations play an important role in stream segregation, our ability to perceive multiple distinct auditory units (Bendixen, Denham, Gyimesi, & Winkler, 2010; Iverson, 1995; Moore & Gockel, 2002), and masking, when one sound renders another inaudible (Buus, 1985). It also is important for recognition of speech (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995) and environmental sounds (Gygi, Kidd, & Watson, 2004), even in the absence of detailed spectral information.

As many natural sounds exhibit great temporal variation, our auditory system often uses this information to aid in understanding our environments. For example, temporal properties are important in modulating acoustic startle responses (ASR) in humans, a low-level reflexive response to acoustic stimuli. As a sound's rise time, or onset, increases, the amplitude and probability of ASR decreases (Blumenthal, 1988). Biases towards specific envelopes may reflect evolutionary pressures, as looming sounds (i.e., those increasing in intensity) are perceived as closer in space to the listener than receding sounds, decreasing intensity (Neuhoff, 2001). This could reflect an adaptive bias towards approaching (e.g., predator, falling tree) vs receding objects.

Amplitude envelope appears to play an integral role in many tasks such as audio-visual integration (Chuen & Schutz, 2016; Grassi & Casco, 2009; Grassi & Pavan, 2012; Schutz & Kubovy, 2009), duration assessment (Grassi & Darwin, 2006; Vallet, Shore, & Schutz, 2014), determining materiality in impact events (Giordano & McAdams, 2006), product preferences, and associative memory (Schutz, Stefanucci, Baum, & Roth, 2017). Exploring its role in both detection in noise and annoyance will add to this growing theoretical understanding of envelope, in addition to applied uses.

Role of Harmonic Structure in Detection and Perception

As spectral structure often receives more research attention than temporal structure (Schutz & Gillard, 2020), it is clear that harmonically 'complex' tones are known to be more detectable than pure tones (Buus, Florentine, & Poulsen, 1997; McPherson, Grace, & McDermott, 2020). Additionally, harmonic sounds (such as those consisting of a fundamental four harmonic components), are more detectable than otherwise equivalent inharmonic tones (McPherson et al., 2020). Suggesting that, advantages in detection may not simply be related to more acoustic energy, but rather the presence and relation of spectral components.

Intensity differences among noise are easier to detect in spectrally invariant noise, compared to spectrally variant random noise (Buus, 1990). Although the complexity of these sounds relative to harmonically invariant tones is closer to natural sounds, they are still far removed from the many sounds we interact with daily. By investigating temporal variation found in environmental sounds, we further our understanding of how we efficiently and effectively interact with our acoustic environment.

Natural sounds exhibit innumerable complexities in spectral content. Musical sounds for example often exhibit temporal shifts in harmonic content, with spectral energy shifting over the course of a tone's duration (Schutz, 2019). These differences in energy concentration are associated with different perceptual centers (p-center) (Howell, 1988). Differences in temporal structure could give rise to very different p-centers even for tones of equal duration, suggesting that p-center captures some aspects of temporal complexity, and it is important for timing of musical (Danielsen et al., 2019) and speech (Marcus, 1981; Morton, Marcus, & Frankish, 1976) sounds. However, its role in tone detection has not been explored.

Present Experiment

As stimuli used in auditory perception research often fail to adequately represent the variety and complexity of sounds in our environment, there is tremendous potential benefit in exploring how acoustic properties common in environmental sounds affect our detection. Through this, we can gain a better understanding of how we detect sound and are able to interact effectively with our acoustic environment. Here, we explore how detectability and perceived annoyance of sounds are affected by changes in their amplitude envelope and harmonic content. As we have not previously explored the effect of harmonic duration, we include the manipulation in order to provide a first step towards understanding whether detectability and annoyance are affected by spectral variation. This manipulation is essentially a variation on 'delayed harmonics', where a tone's fundamental sounds before the harmonics (Edworthy, Loxley, & Dennis, 1991). Additionally, we explore whether changes to a sound's envelope, reducing annoyance, affect their detectability. Based on our team's previous work, (Anderson, 2019; Sreetharan et al., 2021), we hypothesize that percussive sounds will be perceived as less annoying than traditional flat tones.

Methods

Experiment 1

Participants

Eighty-three undergraduate students at a large university in Ontario, Canada enrolled in introductory psychology courses participated. They received course credit in return for their participation. All participants gave informed consent prior to the study in accordance with the university ethics board protocols.

Apparatus

We created the task in *Psychopy* version 2020.1.3 (Peirce et al., 2019), on *Pavlovia* an online study hosting service developed by *Psychopy* (Peirce et al., 2019). Participants completed the experiment using their own computers and headphones.

Stimuli

Our experiment used two types of stimuli (a) target tones and (b) speech noise. We created target tone stimuli in MAESTRO (Ng & Schutz, 2017), a GUI-based sound creation program built on top of the sound synthesis programming language Supercollider (McCartney, 1996). Tones consisted of a 400 Hz fundamental with five harmonics at (800 Hz, 1200 Hz, 1600 Hz, 2000 Hz & 2400 Hz). Percussive and flat tones exhibited a 5ms onset, however percussive tones exponentially decay (for all harmonics), whereas flat tones sustained until 5ms before offset, at which point they decayed. Percussive tones had a total duration of 600ms while flat had a total duration of 360ms, to account for differences in perceived duration. We synthesized based on previous explorations matching assessments of percussive and flat tones (Vallet et al., 2014). To control for differences in perceived loudness between percussive and flat tones, we synthesized each tone with total average root-mean-squared (RMS) of -27.3dB.

Traditional approaches to synthesizing complex sounds (including our team's previous work) synthesize components varying in lock-step with the fundamental— i.e., rising and falling at the same time and in the same proportion across the tone's duration. We varied the length of overtones relative to the fundamental. Specifically, we synthesized the harmonics with five lengths relative to the fundamental's duration: 100%, 50%, 25%, 10%, 5%. We matched amplitudes of all harmonics to the tone's fundamental at onset. Therefore, flat tones exhibited all components in equal amplitudes for either 5%, 10%, etc. of the tone, at which point only the fundamental sounded for the remainder of the tone's nominal duration. For the percussive tones, components decayed with the same offset slope after 5%, 10% etc. of the tone's duration, creating more dynamic variation in the relative mix of the components as the tone sounded.

We used background noise based on speaker-babble from track twenty four of the QuickSIN test, a speech discriminability test designed to test speech comprehension (Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004). The babble consisted of four speakers (one male and three females) speaking English, designed to simulate background speaking in crowded space. We combined the target tone and speech noise stimuli in Amadeus Pro sound editor (HairerSoft, 2019).

Procedure

Consent and Musical Questionnaire.

Participants first gave informed consent, then completed a brief questionnaire on demographics, musical experience, and music listening habits, adapted from the Goldsmith MSI questionnaire (Müllensiefen, Gingras, Stewart, & Musil, 2013).

Experimental Procedure.

After filling in a questionnaire, participants completed a two-part experiment assessing (i) each tone's detectability in noise and (ii) relative annoyance within pairs of tones (drawn from those used in the first task). For the detectability task, participants completed an un-speeded, two interval forced choice (2IFC) tone detection task indicating whether the target tone appeared in interval "1" or "2" of the stimulus. Each interval was indicated by the appearance of a "1" or "2" on the computer screen. We generated sixty unique target tones by factorially combining envelope (percussive or flat), harmonic duration (5%, 10%, 25%, 50%, 100%), signal to noise ratio (4dB, -11dB, -27dB) and interval ("1", or "2"). We presented each target tone twice, resulting in one 120 experimental trials. Participants first heard samples of both the tone and noise used in the experiment, completing two sample trials before the experimental trials.

For the second task (annoyance), participants indicated which of the tones within a pair was most annoying. Participants rated annoyance in two blocks. In the within harmonic block, participants compared tones matched in harmonic durations (e.g., Flat-100% harmonics versus Perc-100% harmonics). Participants made 10 comparisons, with each of the five pairs being presented twice balancing presentation order. In the full comparison block, participants compared two randomly selected tones (10 comparison trials). We randomized the order of the between and random blocks across the participants.

Experiment 2

The within annoyance comparisons in the first experiment showed clearly that percussive tones are less annoying than their harmonic-duration matched counterparts. However, that experiment offered less insight into relative annoyance across different harmonic durations, as the full comparison block included only a randomized subset of all comparisons due to time constraints. To gain a clearer picture of relative annoyance, we conducted a fully balanced annoyance rating experiment pairing each of the 10 tone types with one another. An additional 59 undergraduate students at a large university in Ontario, Canada enrolled in introductory psychology courses participated. They received course credit in return for their participation. All participants gave informed consent prior to the study in accordance with the university ethics board protocols. This experiment consisted of solely of the fully balanced annoyance task. As in experiment 1, we factorially combined two types of envelope (percussive or flat), and five harmonic durations (5%, 10%, 25%, 50%, 100%) resulting in 10 tones. We presented comparisons of each tone with the other nine tones twice, balancing presentation order. In contrast to Experiment 1 which contained 120 signal detection trials and 20 annoyance trials, this experiment contained 90 annoyance trials and no signal detection trials.

Results

Experiment 1

Signal Detection

We used an Analysis of Variance (ANOVA) and Generalized Linear Model (GLM) to independently assess the effects of different sound synthesis approaches on annoyance and detectability. These analyses give novel insight into the relative contributions of amplitude envelope and harmonic duration to these two dimensions of sound important in both improving auditory interface design and enhancing our understanding of basic auditory perception.

Detectability Improves with Percussive Envelopes.

We analyzed the effects of envelope and signal-to-noise ratio (SNR) on tone detection (*Figure 1*) with a two way ANOVA, revealing a significant main effect of envelope, $F(1,9954) = 56.685$, $p < 0.001$, and SNR $F(2, 9954) = 676.360$, $p < 0.001$ on tone detection. Additionally, we found an interaction between these factors $F(2, 9954) = 56.772$, $p < 0.001$. Given the interaction, we explored envelope's simple main effect at each SNR. Although we found no effect of envelope at 4dB SNR, and -11dB SNR, we observed a significant effect at -27dB SNR, $F(1,9954) = 172.8161$, where response accuracy for flat tones (M= 0.56, SD = 0.49) was significantly lower than accuracy for percussive tones (M=0.72, SD = 0.45).

Tukey-HSD post-hoc comparisons on SNR revealed a significant difference between the -27dB (M= 0.64, SD = 0.48) and -11dB (M= 0.91, SD = 0.29) ($p < 0.001$) conditions, as well as between the -27dB (M= 0.64, SD = 0.48) and 4dB (M= 0.93, SD = 0.26) ($p < 0.001$) condition, but not the 4dB and -11dB conditions ($p = 0.1$).

Detectability Affected by Harmonic Duration.

We analyzed the effects of envelope and harmonic condition on tone detection (*Figure 2*) using a two way ANOVA, which found a significant effect of main effect of envelope $F(1,9950) = 49.65 < 0.001$, and harmonic condition $F(4, 9950) = 11.27$, $p < 0.001$. The envelope harmonic interaction was not significant $F(4, 9950) = 1.99$ $p = 0.09$.

Tukey-HSD post-hoc comparisons on the harmonic conditions collapsed across envelope, revealed significantly worse detectability in the 5% harmonic condition (M= 0.78, SD = 0.41) compared to the 25% (M= 0.85, SD = 0.36) ($p < 0.001$), 50% (M= 0.85, SD = 0.36) ($p < 0.001$) and 100% (M= 0.84, SD = 0.37) ($p < 0.001$) condition. As well as between the 10% and 50% condition ($p < 0.05$). We found no other significant differences in detection as a result of harmonic condition, indicating harmonic durations beyond 25% did not significantly improve detectability

Annoyance Affected by Both Envelope and Harmonic Duration

To gain insight into the effect of envelope on annoyance (*Figure 3*), we first evaluated its effect in each of the two conditions using separate chi square tests. In comparison to traditional Flat tones, Percussive tones received significantly lower annoyance ratings in both the 'within' harmonic $X^2(1, N = 83) = 34.005$ $p < 0.0001$, and 'full' comparison conditions $X^2(1, N = 83) = 17.349$, $p < 0.001$. Second, we evaluated the effect of harmonic duration on annoyance in the 'full' condition using a generalized linear model (GLM) with a Poisson link function, using the 5% harmonic as the reference condition for the harmonic variable, and flat as the reference for the envelope variable. In relation to the baseline of the 100% harmonic condition, and adjusting for envelope, rate of annoyance significantly decreased for the 50% (19.3% decrease, $p = 0.03$), 25% (30.7% decrease, $p =< 0.001$), 10% (39.0% decrease, $p =< 0.001$), and 5% (46.9% decrease,

p =< 0.001) conditions. The percussive envelope had a significant decrease (25.2%, $p < 0.001$) of annoyance when compared to flat envelopes, after adjusting for harmonics.

Experiment 2

The second experiment offered each participant the ability to evaluate every comparison, rather than a subset as in the first experiment due to time constraints. We evaluated the effect of harmonic duration and envelope on annoyance (*Figure 4*) with a GLM with a Poisson link function, using the 100% harmonic condition as the reference for the harmonic variable, and percussive as the reference condition for the envelope variable. Consistent with experiment one in relation to the 100% harmonic condition and adjusting for envelope, rate of annoyance significantly decreased for the 50% (12.9% decrease, $p < 0.001$), 25% (21.8% decrease, $p < 0.001$), 10% (39.2% decrease, $p < 0.001$), and 5% (44.6% decrease, $p < 0.001$) conditions. The percussive envelope had a significant decrease in the rate of annoyance ratings when compared to flat envelopes (16.3% decrease, $p < 0.001$), after adjusting for harmonics.

Discussion

Our experiments clarify the role of temporal variation in both detectability and perceived annoyance of auditory stimuli. Offering insight into both our basic understanding of temporal structure in auditory perception, as well as specific guidance for its use in applied settings such as improving auditory interfaces—including auditory alarms in medical devices (Foley, Anderson, & Schutz, 2020). For clarity we will organize this discussion to focus individually on the two main manipulations amplitude envelope and harmonic duration, discussing the implications of these results for both basic and applied contexts.

Amplitude Envelope

One question motivating this project is whether using percussive tones to reduce annoyance potentially reduces signal detection in auditory interfaces. These results suggest it is possible to design interfaces with less annoying percussive sounds that are actually *better detected* in some circumstances. Although further research is needed, we suspect increased detection is due in part to the greater temporal fluctuations in percussive compared to flat tones offering increased contrast and segregation between target and noise.

This result is consistent with work showing the complexities and idiosyncrasies in the temporal structure of musical sounds help segregate sounds (Iverson, 1995), where acoustic parameters such as attack time and the shape of attack are known to aide in differentiation. Listeners are also better at differentiating sounds that are similar in spectral energy when they possess differing envelopes, compared to similar envelopes (Cusack & Roberts, 2000). Together, these investigations show the improvement in perceptual differentiation of sounds by envelope. This differentiation is also advantageous in reducing masking (Buus, 1985), where temporal fluctuations in envelope are less effective at masking, compared to invariant envelope maskers.

Connection to Literature

Literature on acoustic startle response (ASR) also provides some useful complementary context for the superior detection of percussive tones at low listening levels. Temporal parameters such as attack time can affect the probability, amplitude, and latency of ASR. These effects, however, are most evident with low stimulus intensities (Blumenthal, 1988). Therefore, our observed advantage in detection may be indicative of heightened sensitivity to temporal variability at low levels.

Greater sensitivity to temporal changes has interesting connections to a body of work on differential sensitivity to so-called rising vs. falling tones. For example, compared to falling sounds, changes in the intensity (Neuhoff, 1998) and perceived distance from rising tones (Neuhoff, 2001) are over estimated. Although our detection advantage may be indicative of similar perceptual biases to these sounds, rising and falling tones are generally longer in duration than percussive sounds. This difference in duration results in the perception of a sustained source either approaching or receding for rising and falling tones, respectively (Schutz, 2016), while percussive tones are created by (and are generally perceived as) a single impact event. Future research is needed to clarify the degree to which underlying differences in rising and falling tone perception pertain to percussive versus sustained sounds.

Not only do our findings that percussive tones are more easily detected than flat complement and extend previous theoretical work, they hold important practical applications for improving auditory-interfaces such as auditory alarms, which commonly use flat tones (Foley & Schutz, 2021) and where annoyance remains a stubbornly common problem despite decades of research (Block et al., 1999; Gaver, 1997). Reducing masking and increasing perceived differentiation of sounds is imperative as many alarms are temporally invariant it is likely a large contributor to masking (Rayo et al., 2019) and alarm confusions (Edworthy et al., 2017), exacerbated by high alarm rates reducing their efficacy and increasing safety risk for users.

Harmonic Content

Our second question explored the role of harmonic duration in detection as well as annoyance. Spectrally complex tones are detected more readily than pure tones (Buus, Schorer, Florentine, & Zwicker, 1986; McPherson et al., 2020). These sounds, and those used in many alarm standards, typically are temporally invariant, leaving important questions about how temporal changes in spectral structure might be used to improve alerting sounds in devices. In particular, we are interested in whether brief "splashes" of harmonics could achieve some of the detectability benefits of traditional harmonics (i.e. those lasting for the tone's full duration) without some of the problems regarding masking and annoyance inherent with their use.

Work on spectral complexity has shown that listeners better discriminate loudness differences in spectrally invariant noise compared to variant (Buus, 1990), meaning spectrally variant sounds need a larger change in loudness to be perceived as louder or softer. Investigations on recognition of environmental noises, which exhibit large amounts of spectral variation, have shown the vast spectral variability is not required for accurate recognition (Gygi et al., 2004). Although these sounds exhibit great temporal complexity in spectral information, the role of varying spectral information is less understood, possibly resulting in less use with many auditory interfaces.

We found the 25% harmonic condition a potentially beneficial variation on traditional approaches with temporally invariant sounds. Although some decreases in harmonic duration beyond 25% reduced detection, we found no significant increase in detection with longer harmonic durations. In general, participants rated tones with longer durations of harmonic content as more annoying. This suggests using a 25% tone would achieve the same amount of detectability as a 100% tone, but with significantly reduced user annoyance.

To the best of our knowledge, previous investigations have not explored shortening the durations of harmonic content to end earlier than the fundamental. The best context for this work is that of delaying the onset of a sound's harmonics, which reduces its perceived urgency (Edworthy et al., 1991). It should be noted, however, that annoyance and urgency do not perfectly correlate (Marshall et al., 2007). Although further research using a range of temporally variable tones is needed, our findings have important implications for auditory interface design, where such changes, in tandem with envelope, could greatly reduce persistent issues with annoyance (Sreetharan et al., 2021) and masking (Rayo et al., 2019).

Limitations and Future Directions

Due to the coronavirus pandemic, all data collection was completed online, which may have added more noise than traditional in-lab testing with high control of the setting. At the same time, assessment in different contexts is possibly an asset given the potential applied uses of this work. As alarm implementations inevitably differ, ensuring design principles hold true outside of acoustically pristine labs is important.

Our investigation does not clarify why the differences between flat and percussive envelopes are clearest at lower SNR. It may be due to higher peak amplitude of percussive tones. However as changes in envelope have previously been recognized as affecting stream segregation (Cusack & Roberts, 2000; Iverson, 1995), and resistance to masking (Buus, 1985; Hall & Grose, 1988). Further experimenting with tones equated for different temporal envelopes (systematically matching tones for equivalence in peak amplitude vs. total energy vs. perceived loudness) will be important for gaining further insight.

Conclusion

Our data suggest interesting new considerations for constructing auditory interfaces that maximize detectability and minimize annoyance, and add to a growing literature showing differenced in the perception of percussive tones. The detection advantage is consistent with findings that temporal variation aids in segregation (Bendixen et al., 2010; Iverson, 1995; Moore & Gockel, 2002), prevents masking (Buus, 1985), and affects low level startle responses (Blumenthal, 1988). As temporal variation is common in natural sounds, using temporally variable stimuli in our basic perception work will allow for a more complete understanding of our hearing abilities. Through this we can in turn inform and improve auditory interfaces, which currently, paralleling much of basic auditory research, generally fixate on invariant flat tones (Schutz & Gillard, 2020).

From our results, it is clear that design decisions that lead to problems with alarm annoyance are not required (or necessarily helpful) from a detection perspective. These findings indicate two exciting ways in which temporal variation can improve alarm sounds *(a)* percussive tone appear to offer benefits over standard flat tones for both annoyance and detectability, and *(b)* "splashes" of harmonic energy hold potential reducing annoyance and potentially masking, without harming detectability. It appears that relative to proto-typical (100%-Flat) tones our 25%-Percussive tone is considerably less annoying, yet better detected. The benefits are particularly impactful, given that this modification is easy to implement and alarm are prevalent in many safety critical areas such as railways (Edworthy et al., 2011), aviation (Bliss, 2003; Dehais, Roy, Gateau, & Scannella, 2016), the automotive industry (Marshall et al., 2007), industrial processing (Laberge et al., 2014), and healthcare (Edworthy, 2013). Therefore, even small improvements can lead to widespread implications, making many auditory interfaces more pleasant and effective.

References

Anderson, C. J. (2019). *Examining Melodic Alarm Recognition Accuracy Across Contrasting Timbres*.[Unpublished honours thesis]. McMaster University.

Bendixen, A., Denham, S. L., Gyimesi, K., & Winkler, I. (2010). Regular patterns stabilize auditory streams. *The Journal of the Acoustical Society of America*, *128*(6), 3658–3666.

Bliss, J. P. (2003). Investigation of alarm-related accidents and incidents in aviation. *International Journal of Aviation Psychology*, *13*(3), 249–268. https://doi.org/10.1207/S15327108IJAP1303_04

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behaviour in response to alarms of varying reliability. *Ergonomics*, *38*(11), 2300–2312.

Block, F. E., Nuutinen, L., & Ballast, B. (1999). Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance. *Journal of Clinical Monitoring and Computing*, *15*(2), 75–83.

Blumenthal, T. D. (1988). The Startle Response to Acoustic Stimuli Near Startle Threshold: Effects of Stimulus Rise and Fall Time, Duration, and Intensity. *Psychophysiology*, *25*(5), 607–611. https://doi.org/10.1111/j.1469-8986.1988.tb01897.x

Buus, S. (1985). Release from masking caused by envelope fluctuations. *Journal of the Acoustical Society of America*, *78*(6), 1958–1965. https://doi.org/10.1121/1.392652

Buus, S. (1990). Level discrimination of frozen and random noise. *Journal of the Acoustical Society of America*, *87*(6), 2643–2654. https://doi.org/10.1121/1.399057

Buus, S., Florentine, M., & Poulsen, T. (1997). Temporal integration of loudness, loudness discrimination, and the form of the loudness function. *The Journal of the Acoustical Society of America*, *101*(2), 669–680.

Buus, S., Schorer, E., Florentine, M., & Zwicker, E. (1986). Decision rules in detection of simple and complex tones. *Journal of the Acoustical Society of America*, *80*(6), 1646–1657. https://doi.org/10.1121/1.394329

Chuen, L., & Schutz, M. (2016). The unity assumption facilitates cross-modal binding of musical, non-speech stimuli: The role of spectral and amplitude cues. *Attention, Perception, & Psychophysics*, *78*(5), 1512–1528. https://doi.org/10.3758/s13414-016-1088-5

Cusack, R., & Roberts, B. (2000). Effects of differences in timbre on sequential grouping. *Perception and Psychophysics*, *62*(5), 1112–1120. https://doi.org/10.3758/BF03212092

Cvach, M. M. (2012). Monitor alarm fatigue: An integrative review. *Biomedical Instrumentation and Technology*, *46*(4), 268–277. https://doi.org/10.2345/0899-8205-46.4.268

Danielsen, A., Nymoen, K., Anderson, E., Câmara, G. S., Langerød, M. T., Thompson, M. R., & London, J. (2019). Where is the beat in that note? Effects of attack, duration, and frequency on the perceived timing of musical and quasi-musical sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(3), 402–418. https://doi.org/10.1037/xhp0000611

Dehais, F., Roy, R. N., Gateau, T., & Scannella, S. (2016). Auditory alarm misperception in the cockpit: An EEG study of inattentional deafness. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9743*, 177–187. https://doi.org/10.1007/978-3-319-39955-3_17

Edworthy, J. (2013). Medical audible alarms: A review. *Journal of the American Medical Informatics Association*, *20*(3), 584–589. https://doi.org/10.1136/amiajnl-2012-001061

Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *Quality and Safety in Health Care*, *14*(3), 212–215. https://doi.org/10.1136/qshc.2004.013052

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. https://doi.org/10.1016/j.ergon.2010.12.004

Edworthy, J., Loxley, S., & Dennis, I. (1991). Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *Human Factors*, *33*(2), 205–231. https://doi.org/10.1177/001872089103300206

Edworthy, J., Reid, S., McDougall, S., Edworthy, J., Hall, S., Bennett, D., … Pye, E. (2017). The Recognizability and Localizability of Auditory Alarms: Setting Global Medical Device Standards. *Human Factors*, *59*(7), 1108–1127. https://doi.org/10.1177/0018720817712004

Foley, L., Anderson, C. J., & Schutz, M. (2020). Re-Sounding Alarms: Designing Ergonomic Auditory Interfaces by Embracing Musical Insights. *Healthcare*, *8*(4), 389. https://doi.org/10.3390/healthcare8040389

Foley, L., & Schutz, M. (2021). High Time For Temporal Variation: Improving Sonic Interaction With Auditory Interfaces. *IEEE Signal Processing Magazine*.

Gaver, W. W. (1997). Auditory Interfaces. *Handbook of Human-Computer Interaction*, 1003–1041. https://doi.org/10.1016/b978-044481862-1.50108-4

Giordano, B. L., & McAdams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, *119*(2), 1171–1181. https://doi.org/10.1121/1.2149839

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*(6), 1382–1392.

Grassi, M., & Pavan, A. (2012). The subjective duration of audiovisual looming and receding stimuli. *Attention, Perception, & Psychophysics*, *74*(6), 1321–1333. https://doi.org/10.3758/s13414-012-0324-x

Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *The Journal of the Acoustical Society of America*, *115*(3), 1252–1265.

HairerSoft. (2019). Amadeus Pro. Retrieved from https://hairersoft.com/pro.html#1

Hall, J. W., & Grose, J. H. (1988). Comodulation Masking Release: Evidence For Multiple Cues. *Journal of the Acoustical Society of America*, *84*(5), 1669–1675. https://doi.org/10.1121/1.397182

Howell, P. (1988). Prediction of P-center location from the distribution of energy in the amplitude envelope: I. *Perception & Psychophysics*, *43*(1), 90–93. https://doi.org/https://doi.org/10.3758/BF03208978

Iverson, P. (1995). Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 751–763. https://doi.org/http://dx.doi.org/10.1037/0096-1523.21.4.751

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *116*(4), 2395–2405. https://doi.org/10.1121/1.1784440

Laberge, J. C., Bullemer, P., Tolsma, M., & Reising, D. V. C. (2014). Addressing alarm flood situations in the process industries through alarm summary display design and alarm response strategy. *International Journal of Industrial Ergonomics*, *44*(3), 395–406. https://doi.org/10.1016/j.ergon.2013.11.008

Marcus, S. M. (1981). Acoustic determinants of perceptual center (P-center) location. *Perception & Psychophysics*, *30*(3), 247–256. https://doi.org/10.3758/BF03214280

Marshall, D. C., Lee, J. D., & Austria, P. A. (2007). Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. *Human Factors*, *49*(1), 145–157. https://doi.org/10.1518/001872007779598145

McCartney, J. (1996). SuperCollider, a New Real Time Synthesis Language. *ICMC*. Retrieved from https://www.semanticscholar.org/paper/SuperCollider%2C-a-New-Real-Time-Synthesis-Language-McCartney/5f4c5a25c4547f12752e45ec1c692ff740834606

McPherson, M. J., Grace, R. C., & McDermott, J. H. (2020). Harmonicity aids hearing in noise. *BioRxiv*, 2020.09.30.321000. Retrieved from https://doi.org/10.1101/2020.09.30.321000

Moore, B. C. J., & Gockel, H. E. (2002). Factors Influencing Sequential Stream Segregation, *88*(November 2001), 320–332.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, *83*(5), 405–408. https://doi.org/10.1037/0033-295X.83.5.405

Müllensiefen, D., Gingras, B., Stewart, L., & Musil, J. J. (2013). Goldsmiths Musical Sophistication Index (Gold-MSI) v1.0 Technical Report and Documentation Revision 0.3. *Technical Report*, 1–69.

Neuhoff, J. G. (1998). Perceptual bias for rising tones. *Nature*, *395*(6698), 123–124. https://doi.org/10.1038/25862

Neuhoff, J. G. (2001). An adaptive bias in the perception of looming auditory motion. *Ecological Psychology*, *13*(2), 87–110. https://doi.org/10.1207/S15326969ECO1302_2

Ng, M., & Schutz, M. (2017). Seeing sound: A new tool for teaching music perception principles. *Canadian Acoustics*, *45*(3). Retrieved from https://jcaa.caa-aca.ca/index.php/jcaa/article/view/3086

O'Carroll, T. M. (1986). Survey of alarms in an intensive therapy unit. *Anaesthesia*, *41*(7), 742–744. https://doi.org/10.1111/j.1365-2044.1986.tb12844.x

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Rayo, M. F., Patterson, E. S., Abdel-Rasoul, M., & Moffatt-Bruce, S. D. (2019). Using timbre to improve performance of larger auditory alarm sets. *Ergonomics*, *62*(12), 1617–1629. https://doi.org/10.1080/00140139.2019.1676473

Schlesinger, J. J., Baum Miller, S. H., Nash, K., Bruce, M., Ashmead, D. H., Shotwell, M. S., … Weinger, M. B. (2018). Acoustic features of auditory medical alarms—An experimental study of alarm volume. *The Journal of the Acoustical Society of America*, *143*(6), 3688–3697. https://doi.org/10.1121/1.5043396

Schutz, M. (2016). Clarifying amplitude envelope's crucial role in auditory perception. *Canadian Acoustics*, *44*(2), 42–43.

Schutz, M. (2019). Acoustic structure and musical function: Musical notes informing auditory research. In M. H. Thaut & D. A. Hodges (Eds.), *The Oxford Handbook on Music and the Brain* (pp. 145–164). Oxford, UK: Oxford University Press.

Schutz, M., & Gillard, J. (2020). On the generalization of tones: A detailed exploration of non-speech auditory perception stimuli. *Scientific Reports*, *10*(1), 9520. https://doi.org/10.1038/s41598-020-63132-2

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., Baum, S. H., & Roth, A. (2017). Name that percussive tune: Associative memory and amplitude envelope. *The Quarterly Journal of Experimental Psychology*, *70*(7), 1323–1343. https://doi.org/http://dx.doi.org/10.1080/17470218.2016.1182562

Sendelbach, S., & Funk, M. (2013). Alarm fatigue: A patient safety concern. *AACN Advanced Critical Care*, *24*(4), 378–386. https://doi.org/10.1097/NCI.0b013e3182a903f9

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303–304. https://doi.org/10.1126/science.270.5234.303

Sreetharan, S., Schlesinger, J. J., & Schutz, M. (2021). Decaying amplitude envelopes reduce alarm annoyance: Exploring new approaches to improving auditory interfaces. *Applied Ergonomics*, *96*.

Vallet, G., Shore, D. I., & Schutz, M. (2014). Exploring the role of amplitude envelope in duration estimation. *Perception*, *43*(7), 616–630.

Varpio, L., Kuziemsky, C., MacDonald, C., & King, W. J. (2012). The helpful or hindering effects of in-hospital patient monitor alarms on nurses: A qualitative analysis. *CIN - Computers Informatics Nursing*, *30*(4), 210–217. https://doi.org/10.1097/NCN.0b013e31823eb581

Figures

Figure 1. Proportion of correct responses in signal detection task by envelope and signal to noise ratio. Error bars denote standard error.

Figure 1.

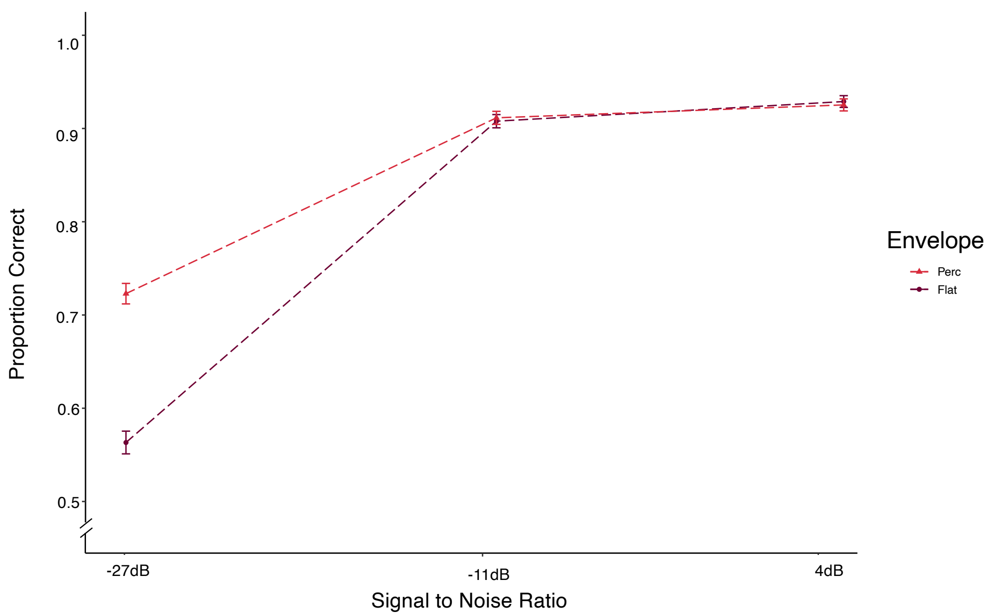


Figure 2. Proportion of correct responses in signal detection task by envelope and harmonic duration condition. Error bars denote standard error.

Figure 2.

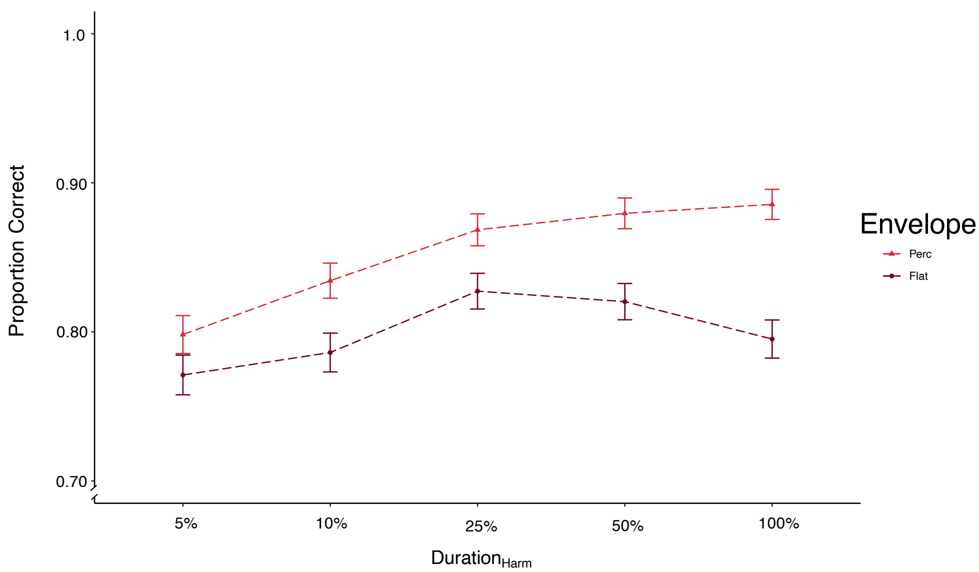


Figure 3. Annoyance ratings across 'Within Harmonic Comparisons' and 'Full Comparisons'.

Figure 3.

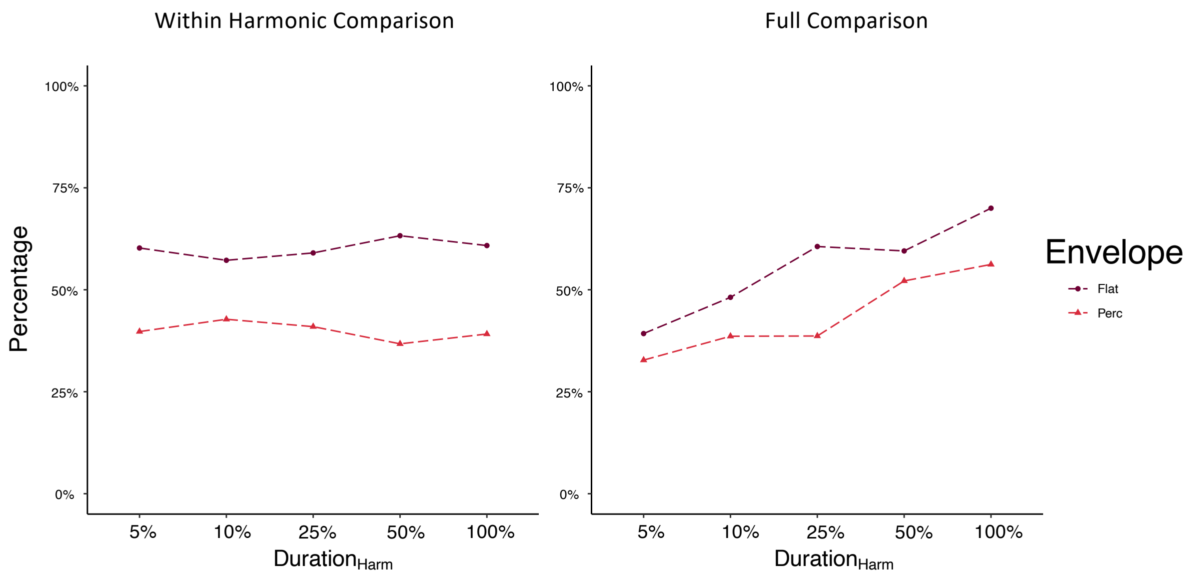
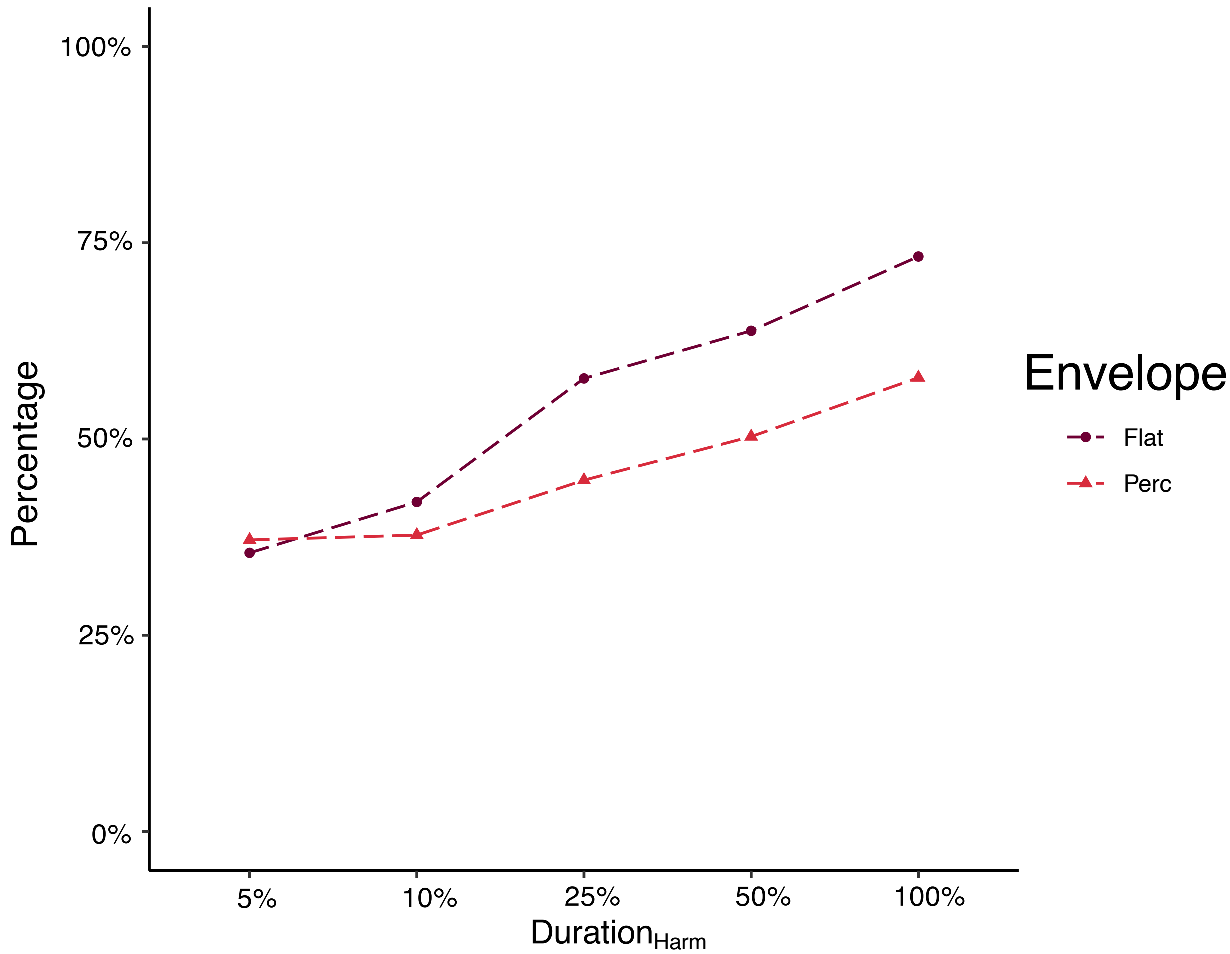


Figure 4. Annoyance ratings in the fully balanced condition.

Figure 4.



Chapter 3

Improving Detectability of Auditory Interfaces Through Temporal Variation in Envelope

Liam Foley, Joseph Schlesinger, Michael Schutz

Abstract

Our sonic interactions are increasingly with designed sounds from auditory interfaces, including auditory alarms. These sounds play a vital role in many industries alerting users to system and safety critical information. Unfortunately, problems such as high alarm rate, low reliability, and annoyance reduce their efficacy. Here we explore how to reduce annoyance and improve detection by manipulating a sound's temporal envelope. In the first experiment participants completed a tone detection task concurrently with a speech comprehension task. Tones were either a temporally variant percussive tone or an invariant flat tone presented at six signal-to-noise ratios (4dB, -11dB, -18dB, -21dB, -24dB, -27dB). In a second experiment participants completed an annoyance rating task with the tones used in experiment one. Participants detected percussive tones more accurately than flat tones, especially at low signal levels, and rated percussive tones as less annoying than flat tones. These results provide more evidence that temporal variation of amplitude envelope is a promising design parameter for future alarms.

Introduction

Auditory interfaces help us interact with numerous devices. One such interface – auditory alarms – are particularly important in safety-critical areas as they offer faster response times over visual alerts (Morris & Montano, 1996) and are widespread through high-consequence industries including industrial processing (Laberge, Bullemer, Tolsma, & Reising, 2014), navigation systems in aviation (Bliss, 2003), automobiles, (Marshall, Lee, & Austria, 2007), railways (Edworthy, Hellier, Titchener, Naweed, & Roels, 2011) and healthcare (Edworthy, 2013).

Although auditory alerts are of high value to enhancing human-computer interactions, several key problems still preclude realization of their full potential. A common challenge in healthcare (Varpio, Kuziemsky, MacDonald, & King, 2012) and industrial processing (Laberge et al., 2014) is that of high alarm rate. In part, this reflects a "better safe than sorry" design philosophy, where alarms are set to sound often to ensure they are heard (Edworthy & Hellier, 2005; Patterson, 1990). Studies of healthcare alarms have shown rates of up to more than three hundred and fifty alarms per patient per day (Sendelbach & Funk, 2013). Although these alerts can be informative, rarely do they indicate urgent action is required — for example, one report found less than 0.5% (i.e. eight out of 1455) indicate a life threatening event (O'Carroll, 1986).

The combination of high alarm rates and with low reliability, directly affects users' feelings of annoyance (Lerner, Dekker, Steinberg, & Huey, 1996) and creates an incentive for tuning them out or even turning them off—potentially resulting in disastrous consequences (Bliss, Gilson, & Deaton, 1995; Block, Nuutinen, & Ballast, 1999). For these reasons, many alarm modifications have focused on reducing the number of alarms to improve alarm efficacy. Here we investigate ways to improve alarm efficacy by modifying the sounds themselves.

Alarm Modifications and Interventions

Numerous interventions have been used in attempt to improve alarm efficacy. In healthcare, interventions include adding secondary notifications to staff pagers (Cvach, Frank, Doyle, & Stevens, 2014), changing patient electrode placement procedures (Leigher, Kemppainen, & Neyens, 2020) and customizing alarm settings for low-risk patients (Dandoy et al., 2014) in an attempt to reduce the number of alarms. Although these interventions have shown some promise in reducing alarm rates and resulting fatigue, there has been very little attention on *improving the sounds themselves*. Starting with better designed and perceptually validated sounds alarm modifications go from reactive interventions, attempting to fix problems, to alarms designed to prevent these issues. As alarms rates are likely to increase as devices proliferate (Borowski et al., 2011)**,** improvements to baseline best practices for sound design will ensure far safer and more usable alarm systems.

Study Motivation

Here we ask whether changes to an alarm sound's amplitude envelope (i.e., the way amplitude changes over time) can reduce annoyance while maintaining detectability. This acoustic property has been under explored in both alarm design (Foley, Anderson, & Schutz, 2020) and auditory perception (Schutz & Gillard, 2020), yet it is known to play an important role in perceptual abilities including audio-visual integration (Chuen & Schutz, 2016; Grassi & Casco, 2009; Grassi & Pavan, 2012; Schutz & Kubovy, 2009), duration perception (Grassi & Darwin, 2006; Vallet, Shore, & Schutz, 2014), and determining materiality from impact events (Giordano & McAdams, 2006). The results of previous investigations also suggest that manipulating a sound's amplitude envelope may result in better auditory alarms. For example, sounds with temporally variant envelopes are better associated with items in associative memory

(Schutz, Stefanucci, Baum, & Roth, 2017), detected better at low intensity levels, (Chapter 2) and are rated as less annoying than tones with invariant envelopes (Chapter 2; Sreetharan, Schlesinger, & Schutz, 2021).

“Annoyance” is a complicated construct to measure, with numerous factors recognized as playing a role. For example, musical preference is known to modulate both perceived intensity and annoyance of musical stimuli (Fucci, Petrosino, Hallowell, Andra, & Wilcox, 1997). Participants subjective view of the “appropriateness” of an alarm tone can also influence annoyance ratings (Lerner et al., 1996). Perceived loudness and detectability of sound is also correlated with annoyance (Berglund, Preis, & Rankin, 1990; Fidell & Teffeteller, 1981; Fucci et al., 1997; Sneddon, Howe, Pearsons, & Fidell, 1996; Steele & Chon, 2007). Severl studies have found that temporally variant sounds are much less annoying than conventional invariant sounds (Anderson, 2019; Chapter 2; Sreetharan et al., 2021). Importantly reductions in annoyance did not come at the cost of reduced learning or detectability of the sounds.

Present Study

In the present study we seek to further explore the role temporal variation in envelope plays in detectability, while completing a secondary task, and annoyance of sounds. Here we use a concurrent speech task adapted from one previously used in a study on alarm volumes (Schlesinger et al., 2018) to simulate alarm environments. We hypothesize that percussive sounds will be detected better at lower signal to noise ratios (SNR) and perceived as less annoying than flat sounds. In addition to lending insight into a the fundamental process of tone detection, it holds important applications for ongoing efforts to design more effective sounds in auditory devices and for clarifying the role envelope plays in our auditory abilities.

Methods

Experiment 1

Participants

Sixty undergraduate students enrolled in introductory psychology courses participated in the current experiment. They received course credit in return for their participation. All participants gave informed consent prior to the study in accordance with the university ethics board protocols, and they received course credit in return for their participation.

Apparatus

We created the task in *Psychopy* version 2020.1.3 (Peirce et al., 2019), on *Pavlovia* an online study hosting service developed by *Psychopy* (Peirce et al., 2019). Participants completed the experiment using their own computers and headphones.

Stimuli

The experiment consisted of two types of stimuli: (a) target tones and (b) speaker phrases. We created target tones in MAESTRO (Ng & Schutz, 2017), a GUI-based sound creation program built on top of the sound synthesis programming language Supercollider (McCartney, 1996). Tones consisted of a 400Hz fundamental with five harmonics (800hz, 1200hz, 1600Hz, 2000Hz, 2400Hz). Flat tones exhibited a relatively invariant amplitude envelope that consisted of a 5ms onset, an invariant sustain, and a 5ms offset. Percussive tones exhibited a variant amplitude envelope that consisted of a 5ms onset, followed by an exponential decay. Flat tones had a total duration of 360ms while percussive tones had a total duration of 600ms, in order to account for differences in perceived duration (Vallet et al., 2014). We generated twelve target tone stimuli fully factorial across envelope (flat & percussive), and signal-to-noise ratios (SNR; 4dB, -11dB, -18dB, -21dB,-24dB,-27dB).

We used a standardized set of speech phrases from the coordinated response measure (CRM) task for speech intelligibility assessment in multichannel environments (Bolia et al., 2000). All phrases consisted of the following structure "Ready (call sign), go to (colour) (number) now". We created 216 combinations for our target phrase using the call sign "baron", four colours (red, green, blue, white) and four numbers (1,2,3,& 4). To create distractor phrases we randomly choose two phrases from callsigns (Arrow, Charlie, Eagle, Hopper, Laker, Ringo, Tiger), colour (red, green, blue, white), number (1,2,3,4) combinations.

Procedure

Consent and Musical Questionnaire.

Participants first gave informed consent, then completed a brief questionnaire on demographics, musical experience, and music listening habits, adapted from the Goldsmith MSI questionnaire (Müllensiefen, Gingras, Stewart, & Musil, 2013).

Experimental Procedure.

Participants concurrently completed two tasks (i) a speech comprehension task and a (ii) signal detection task. The speech comprehension task used three CRM phrases a target phrase and two distractors. CRM phrases were played 500 ms apart, with the first distractor phrase, followed by the target phrase, and then the second distractor phrase. Participants indicated what colour – number was spoken by the Baron speaker by clicking on screen. For the signal detection task participants indicated when they heard a tone by pressing the space bar. The tone appeared on half of the speech comprehension trails at one of three possible tone onsets (750ms, 1250ms, 2000ms) for a total of 216 experimental trials.

Experiment 2

Experiment two consisted of paired annoyance ratings of the target tones used in experiment one. It did not include any detection or speech comprehension task. Since the SNR of each tone was relative to the speaker babble in experiment one these tone conditions are relative loudness levels. The average root mean squared (RMS) level of each tone was -22dB (4dB SNR), -37dB (-11dB SNR), -44dB (-18dB SNR), -47dB (-21dB SNR), -50dB (-24dB SNR), and -53dB (-27dB SNR).

Participants

Thirty-five participants recruited from Prolific, an online recruitment service (Prolific, 2021), completed the experiment. They received compensation of 7.50 GBP an hour prorated, as per Prolific's compensation guidelines. All participants gave informed consent prior to the study in accordance with the university ethics board protocols.

Procedure

Participants completed the same consent and questionnaire as in experiment one. Using the target tones used in experiment one, participants indicated which of a pair of tones they found the most annoying. We presented all envelope and RMS combinations counterbalanced for presentation order resulting in 132 trials.

Results

Experiment 1

Signal Detection

We calculated *d'* (Macmillan & Creelman, 2005) for each participant in each tone condition (flat, percussive), at each SNR (4dB,-11dB,-18dB,-21dB,-24dB,-27dB). These results are shown in Figure 1. We analyzed the effects of envelope and with a two-way ANOVA. This revealed a significant main effect of envelope $F(1,708) = 6.35$, $p < 0.05$, and a fallow-up t-test found that d' was higher for percussive ($M = 1.59, SD = 0.99$) tones compared to flat ($M = 1.42, SD = 0.89$), $t(710.34) = -2.45$, $p < 0.05$. We also found a main effect of SNR, $F(5,708) = 9.03$, $p < 0.05$, reflecting an expected relationship between sound level and detectability. The two-way interaction was not significant, $F(5,708) = 0.69$, $p = 0.62$, suggesting the envelope's effect does not vary significantly across SNR. We further explored envelope by analyzing the simple main effect of envelope at each SNR. We found that d' was significantly higher for percussive tones (M=1.33, SD = 0.86) than flat tones (M = 0.97, SD=0.55) at the -27 dB SNR condition but not in others $F(1,708) = 4.43$, $p < 0.05$, but not at the other SNRs.

CRM Performance

Performance in the speech task is plotted in Figure 2. We analyzed speech task performance in the different target tone conditions with a two way ANOVA. This revealed a non-significant main effect of envelope $F(1,12947)= 2.03$, $p > 0.05$, a significant main effect of SNR $F(6,12947)= 3.07$, $p < 0.05$, indicating the level of the tone signal affected participants performance in the speech task. Additionally, we found a significant interaction between SNR and envelope $F(5,12947)=2.65$, $p < 0.05$.

Experiment 2

The results of experiment two are shown in Figure 3. To evaluate tone annoyance, we first assessed the effect of SNR using a generalized linear model with a Poisson link function. Across SNR levels, percussive tones are rated significantly less annoying than flat tones (-23.7%, $p < 0.001$). Across envelopes, annoyance in all conditions is significantly lower than the -22dB condition: -37dB (-31.2%, $p < 0.001$), -44dB(-47.2%, $p < 0.001$), -47dB (-51.2%, $p < 0.001$), -50dB (-58.0%, $p < 0.001$), and -53dB (-61.5%, $p < 0.001$) conditions.

Discussion

Here we show that in relation to standard flat tones, percussive tones offer a reduction in annoyance paired with improved detection accuracy—without harming comprehension in a concurrent speech task. Together, these findings hold tremendous potential for improving the design of auditory alarms. This is crucial insight extending current approaches, as existing efforts to mitigate alarm issues rarely explore the alarm sounds themselves. Therefore, our finding that simple changes in acoustic properties can improve their efficacy (while retaining any pitch/timing patterns currently used) holds great promise and applicability.

Tone Detection Task

Amplitude Envelope

In previous investigations participants detected percussive tones significantly better than flat tones when solely focused on detection (Chapter 2). One important consideration for applied purposes is whether these detection benefits translate to situations when attention is divided into two or more tasks. In many real world uses of alarms (e.g. hospitals), staff need to attend to a variety of auditory signals, including speech, concurrent with alarms. It is critical that new sounds do not impair effective speech communication as well as maintain signal detectability. Therefore our results showing that temporal variation enhances tone detection while preserving speech comprehension and maximizing ergonomic utility (i.e., minimizing annoyance) further strengthens the case for exploring a wider palate of temporally variable sounds in auditory interface design.

Signal-to-Noise Ratio

Our results indicating alarm sounds do not need to greatly exceed background noise levels to be detectable are consistent with previous findings (Chapter 2; Schlesinger et al., 2018), but are contrary to conventional design beliefs that louder is more detectable and thus better. Our results suggest that it may be time to re-think practices in healthcare environments, where overall noise levels are often above recommended safe levels (Oleksy & Schlesinger, 2019). As alarm rate is very high in many contexts (e.g., hospitals, airline cockpits, nuclear power plants), we believe the option to manipulate the temporal structure of alarms to reduce annoyance and preserve detectability at lower SNRs holds significant potential for improving the auditory environment of many workplaces.

CRM Task Performance

Performance in the CRM task is consistent with previous findings showing a trade-off between performance in the speech comprehension and detection tasks (Schlesinger et al., 2018). Here, louder tones (i.e., the higher SNR condition) lead to increased accuracy in the detection task but decreased accuracy on the CRM, whereas quieter tones lead to decreased detection but increased CRM accuracy. Importantly, tone envelope had no effect on this trade-off, suggesting that the use of more temporally varied tones as alarm sounds would not affect known trade-offs related to signal level. This again suggests it to be a fruitful parameter to explore to improve alarm efficacy without distracting from users' needs for speech communication.

Annoyance Rating Task

Consistent with our previous investigations, we find percussive tones rated as significantly less annoying than flat tones (Chapter 2; Sreetharan et al., 2021). Additionally we see a decrease in annoyance ratings for relatively softer tones, consistent with general findings that relative intensity correlates well with annoyance ratings (Berglund et al., 1990; Fucci et al., 1997; Steele & Chon, 2007).

As many alarm systems have high rates but low reliability, mitigating annoyance will help garner more user acceptance of alarm systems (Lerner et al., 1996). Importantly these reductions in annoyance do not come at the cost of detectability. Despite detectability often correlating with annoyance ratings (Fidell & Teffeteller, 1981; Sneddon et al., 1996), here we provide more evidence that although annoyance is sufficient for detection, it is not necessary.

Connection to Literature

This investigation adds to a growing understanding of how temporal variation can be a useful albeit underexplored acoustic property for improving auditory alarms. Through the use of temporal variation we can create more effective alarms by not only improving detection, but also improving recognition (Gygi, Kidd, & Watson, 2004; Zeng & Shannon, 1995), preventing masking (Buus, 1985; Rayo, Patterson, Abdel-Rasoul, & Moffatt-Bruce, 2019), and maintaining learnability (Sreetharan et al., 2021), which is low in some alarm standards (Sanderson, Wee, & Lacherez, 2006).

Further, the current study adds to our understanding the role temporal variation plays in annoyance and preference of sounds (McDermott, 2012). Although annoyance is correlated with perceived intensity, the association between annoyance and intensity can be modulated (particularly at low levels) by other factors. These include preference of the sound (Fucci et al., 1997), sharpness (Berglund et al., 1990), onset, duration, grouping (Marshall et al., 2007), as well as our findings here regarding amplitude envelope. Future exploration and applications of temporal variation could help create more effective and less obtrusive alarms sounds.

Sound Design as Alarm Intervention

Manipulations of amplitude envelope are trivial from the perspective of sound synthesis. In video games, sounds often exhibit great temporal variation creating acoustic realism to further immerse the user (Salselas, Penha, & Bernardes, 2020). Ironically, sounds often are overly simplistic in applications where sounds are integral to user safety. Although historically technological limitations may have limited the scope of what sounds could be used (Edworthy & Hards, 1999), today's sounds reflect limitations of imagination, rather than technical constraints—which has resulted in an unfortunate situation where many of the most important alarms sound the worst.

The need for better alarms is paramount because the number of alarms is likely to increase in many industries (Borowski et al., 2011). Improving alarms aesthetics is important, as it can directly lead to users belief in their usability and acceptance (Kiefer et al., 1999). Merely dismissing alarms on their aesthetic merits set back usability in medical alarms (Block, 2008), however, we should also be critical of accepting future alarms without perceptually validated backing to avoid new alarms that sound 'better' but are problematic form other perspectives.

Limitations

Due to the covid-19 pandemic this study was completed online and participants completed the experiment using their own computer and headphones. Compared to traditional in-lab testing approaches, our lack of control over equipment likely introduces more variability into the data. However, given the broad implementation of alarms in environments that are not tailored for pristine auditory playback, the lack of control over testing may add a degree of robustness to our results.

Conclusion

Simple manipulations of a sound's envelope can help to address problems with auditory alarms, that are prevalent in many safety critical areas such as medicine (Sanderson, Liu, & Jenkins, 2009), industrial processing (Laberge et al., 2014), aviation (Bliss, 2003; Dehais, Roy, Gateau, & Scannella, 2016), and the automotive industry (Marshall et al., 2007). Rather than recommending a specific sound, we see this work as the basis for larger future inquiries into the role of temporal variation as an invaluable acoustic parameter for auditory interfaces. Our data add to mounting evidence that temporal variation holds is an incredibly valuable design parameter. Particularly in improving detection, while minimising annoyance. Through more perceptually validated sound design, environments can be less noisy, sound better, and most importantly, safer.

References

Anderson, C. J. (2019). *Examining Melodic Alarm Recognition Accuracy Across Contrasting Timbres*.

Berglund, B., Preis, A., & Rankin, K. (1990). Relationship between loudness and annoyance for ten community sounds. *Environment International*, *16*(4–6), 523–531. https://doi.org/10.1016/0160-4120(90)90021-W

Bliss, J. P. (2003). Investigation of alarm-related accidents and incidents in aviation. *International Journal of Aviation Psychology*, *13*(3), 249–268. https://doi.org/10.1207/S15327108IJAP1303_04

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behaviour in response to alarms of varying reliability. *Ergonomics*, *38*(11), 2300–2312.

Block, F. E. (2008). "For if the trumpet give an uncertain sound, who shall prepare himself to the battle?" (I Corinthians 14:8, KJV). *Anesthesia & Analgesia*, *106*(2), 357–359. https://doi.org/10.1213/ane.0b013e3181606927

Block, F. E., Nuutinen, L., & Ballast, B. (1999). Optimization of alarms: A study on alarm limits, alarm sounds, and false alarms, intended to reduce annoyance. *Journal of Clinical Monitoring and Computing*, *15*(2), 75–83.

Borowski, M., Görges, M., Fried, R., Such, O., Wrede, C. E., & Imhoff, M. (2011). Medical device alarms. *Biomedizinische Technik- Biomedical Engineering*, *56*(2), 73–83.

Buus, S. (1985). Release from masking caused by envelope fluctuations. *Journal of the Acoustical Society of America*, *78*(6), 1958–1965. https://doi.org/10.1121/1.392652

Chuen, L., & Schutz, M. (2016). The unity assumption facilitates cross-modal binding of musical, non-speech stimuli: The role of spectral and amplitude cues. *Attention, Perception,*

& Psychophysics, *78*(5), 1512–1528. https://doi.org/10.3758/s13414-016-1088-5

Cvach, M. M., Frank, R. J., Doyle, P., & Stevens, Z. K. (2014). Use of pagers with an alarm escalation system to reduce cardiac monitor alarm signals. *Journal of Nursing Care Quality*, *29*(1), 9–18. https://doi.org/10.1097/NCQ.0b013e3182a61887

Dandoy, C. E., Davies, S. M., Flesch, L., Hayward, M., Koons, C., Coleman, K., … Weiss, B. (2014). A team-based approach to reducing cardiac. *Pediatrics*, *134*(6), 1–11. https://doi.org/10.1542/peds.2014-1162

Dehais, F., Roy, R. N., Gateau, T., & Scannella, S. (2016). Auditory alarm misperception in the cockpit: An EEG study of inattentional deafness. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9743*, 177–187. https://doi.org/10.1007/978-3-319-39955-3_17

Edworthy, J. (2013). Medical audible alarms: A review. *Journal of the American Medical Informatics Association*, *20*(3), 584–589. https://doi.org/10.1136/amiajnl-2012-001061

Edworthy, J., & Hards, R. (1999). Learning auditory warnings: The effects of sound type, verbal labelling and imagery on the identification of alarm sounds. *International Journal of Industrial Ergonomics*, *24*(6), 603–618.

Edworthy, J., & Hellier, E. (2005). Fewer but better auditory alarms will improve patient safety. *Quality and Safety in Health Care*, *14*(3), 212–215. https://doi.org/10.1136/qshc.2004.013052

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. https://doi.org/10.1016/j.ergon.2010.12.004

Fidell, S., & Teffeteller, S. (1981). Scaling The Annoyance of Intrusive Sounds. *Journal of*

Sound and Vibration, *78*(3), 291–298. https://doi.org/0.1016/S0022-460X(81)80039-9

Foley, L., Anderson, C. J., & Schutz, M. (2020). Re-Sounding Alarms: Designing Ergonomic Auditory Interfaces by Embracing Musical Insights. *Healthcare*, *8*(4), 389. https://doi.org/10.3390/healthcare8040389

Foley, L., Schlesinger, J., & Schutz, M. (n.d.). More Detectable, Less Annoying. The Role of Temporal Variation in Envelope and Spectral Content on Detection and Annoyance. *Psychological Science*.

Fucci, D., Petrosino, L., Hallowell, B., Andra, L., & Wilcox, C. (1997). Magnitude estimation scaling of annoyance in responce to rock music: Effects of sex and listenerrs prereference. *Perceptual and Motor Skills*, *84*(2), 663–670.

Giordano, B. L., & McAdams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, *119*(2), 1171–1181. https://doi.org/10.1121/1.2149839

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*(6), 1382–1392.

Grassi, M., & Pavan, A. (2012). The subjective duration of audiovisual looming and receding stimuli. *Attention, Perception, & Psychophysics*, *74*(6), 1321–1333. https://doi.org/10.3758/s13414-012-0324-x

Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *The Journal of the Acoustical Society of America*, *115*(3), 1252–

1265.

Kiefer, R., LeBlanc, D., Palmer, M., Salinger, J., Deering, R., & Shulman, M. (1999). Development and validation of functional definitions and evaluation procedures for collision warning/avoidance systems: Final report. *NHTSA Technical Support*. https://doi.org/10.1037/e495792008-001

Laberge, J. C., Bullemer, P., Tolsma, M., & Reising, D. V. C. (2014). Addressing alarm flood situations in the process industries through alarm summary display design and alarm response strategy. *International Journal of Industrial Ergonomics*, *44*(3), 395–406. https://doi.org/10.1016/j.ergon.2013.11.008

Leigher, D., Kemppainen, P., & Neyens, D. M. (2020). Skin preparation and electrode replacement to reduce alarm fatigue in a community hospital intensive care unit. *American Journal of Critical Care*, *29*(5), 390–395. https://doi.org/10.4037/ajcc2020120

Lerner, N. D., Dekker, D. K., Steinberg, G. V., & Huey, R. W. (1996). *Inappropriate Alarm Rates and Driver Annoyance*. Retrieved from https://rosap.ntl.bts.gov/view/dot/13303

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory : A User's Guide* (Vol. 2nd ed). Mahwah, N.J.: Psychology Press. Retrieved from https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,url,uid&db=e000xna&AN=119239&site=ehost-live&scope=site

Marshall, D. C., Lee, J. D., & Austria, P. A. (2007). Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. *Human Factors*, *49*(1), 145–157. https://doi.org/10.1518/001872007779598145

McCartney, J. (1996). SuperCollider, a New Real Time Synthesis Language. *ICMC*. Retrieved from https://www.semanticscholar.org/paper/SuperCollider%2C-a-New-Real-Time-

Synthesis-Language-McCartney/5f4c5a25c4547f12752e45ec1c692ff740834606

McDermott, J. H. (2012). Auditory preferences and aesthetics: Music, voices, and everyday sounds. In *Neuroscience of Preferance and Choice*. Academic Press. https://doi.org/10.1016/B978-0-12-381431-9.00020-6

Morris, R. W., & Montano, S. R. (1996). Response times to visual and auditory alarms during anaesthesia. *Anaesthesia and Intensive Care*, *24*(6), 682–684. https://doi.org/10.1177/0310057x9602400609

Müllensiefen, D., Gingras, B., Stewart, L., & Musil, J. J. (2013). Goldsmiths Musical Sophistication Index (Gold-MSI) v1.0 Technical Report and Documentation Revision 0.3. *Technical Report*, 1–69.

Ng, M., & Schutz, M. (2017). Seeing sound: A new tool for teaching music perception principles. *Canadian Acoustics*, *45*(3), 104–105. Retrieved from https://jcaa.caa-aca.ca/index.php/jcaa/article/view/3086

O'Carroll, T. M. (1986). Survey of alarms in an intensive therapy unit. *Anaesthesia*, *41*(7), 742–744. https://doi.org/10.1111/j.1365-2044.1986.tb12844.x

Oleksy, A. J., & Schlesinger, J. J. (2019). What's all that noise—Improving the hospital soundscape. *Journal of Clinical Monitoring and Computing*, *33*(4), 557–562. https://doi.org/10.1007/s10877-018-0215-3

Patterson, R. D. (1990). Auditory warning sounds in the work environment. *Philosophical Transactions of the Royal Society of London*, *327*(1241), 485–492.

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., … Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Prolific. (2021). Prolific. Oxford, UK. Retrieved from https://www.prolific.co

Rayo, M. F., Patterson, E. S., Abdel-Rasoul, M., & Moffatt-Bruce, S. D. (2019). Using timbre to improve performance of larger auditory alarm sets. *Ergonomics*, *62*(12), 1617–1629. https://doi.org/10.1080/00140139.2019.1676473

Salselas, I., Penha, R., & Bernardes, G. (2020). Sound design inducing attention in the context of audiovisual immersive enviornments. *Personal and Ubiquitous Computing*, *4*(4), 1–12. https://doi.org/https://doi.org/10.1007/s00779-020-01386-3

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.

Schlesinger, J. J., Baum Miller, S. H., Nash, K., Bruce, M., Ashmead, D. H., Shotwell, M. S., … Weinger, M. B. (2018). Acoustic features of auditory medical alarms—An experimental study of alarm volume. *The Journal of the Acoustical Society of America*, *143*(6), 3688–3697. https://doi.org/10.1121/1.5043396

Schutz, M., & Gillard, J. (2020). On the generalization of tones: A detailed exploration of non-speech auditory perception stimuli. *Scientific Reports*, *10*(1), 9520. https://doi.org/10.1038/s41598-020-63132-2

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., Baum, S. H., & Roth, A. (2017). Name that percussive tune: Associative memory and amplitude envelope. *The Quarterly Journal of Experimental Psychology*, *70*(7), 1323–1343.

https://doi.org/http://dx.doi.org/10.1080/17470218.2016.1182562

Sendelbach, S., & Funk, M. (2013). Alarm fatigue: A patient safety concern. *AACN Advanced Critical Care*, *24*(4), 378–386. https://doi.org/10.1097/NCI.0b013e3182a903f9

Sneddon, M., Howe, R., Pearsons, K., & Fidell, S. (1996). Annoyance Ratio Study of the Noticeability and of Sounds of Low Signal-to-Noise. *NASA Contractor Report*, 1–98.

Sreetharan, S., Schlesinger, J. J., & Schutz, M. (2021). Decaying amplitude envelopes reduce alarm annoyance: Exploring new approaches to improving auditory interfaces. *Applied Ergonomics*, *96*.

Steele, D. L., & Chon, S. H. (2007). A perceptual study of sound annoyance. *Audio Mostly 2007 - 2nd Conference on Interaction with Sound, Conference Proceedings*, (January 2007), 19–24.

Vallet, G., Shore, D. I., & Schutz, M. (2014). Exploring the role of amplitude envelope in duration estimation. *Perception*, *43*(7), 616–630.

Varpio, L., Kuziemsky, C., MacDonald, C., & King, W. J. (2012). The helpful or hindering effects of in-hospital patient monitor alarms on nurses: A qualitative analysis. *CIN - Computers Informatics Nursing*, *30*(4), 210–217. https://doi.org/10.1097/NCN.0b013e31823eb581

Zeng, F.-G., & Shannon, R. V. (1995). Possible origins of the non-monotonic intensity disctimination function in forward masking. *Hearing Research*, *82*(2), 216–224.

Figures

Figure 1. Tone Detection Performance for Flat and Percussive Tones Across SNR. Error bars denote standard error.

Figure 1.

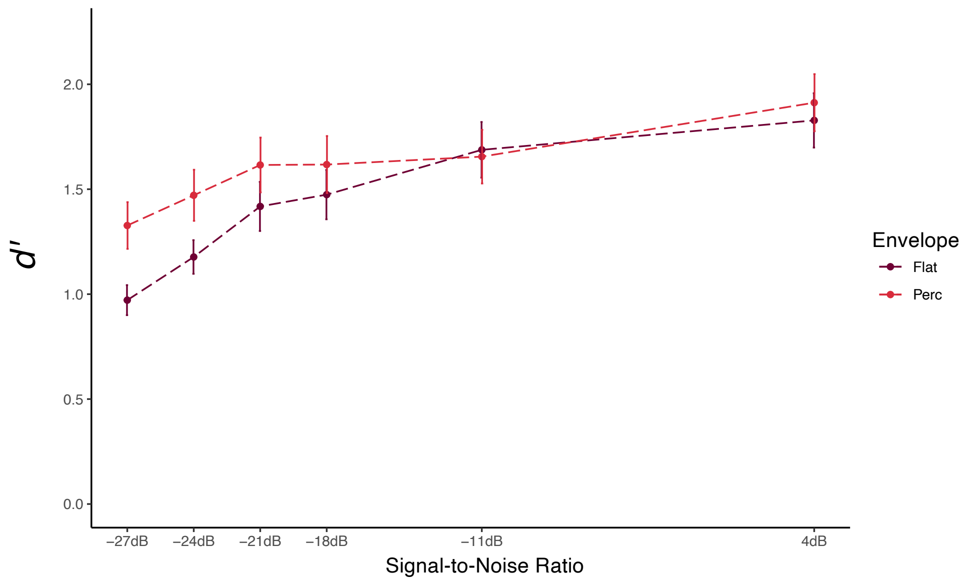


Figure 2. Speech Task Performance Across Tone Detection Conditions. Error bars denote standard error.

Figure 2.

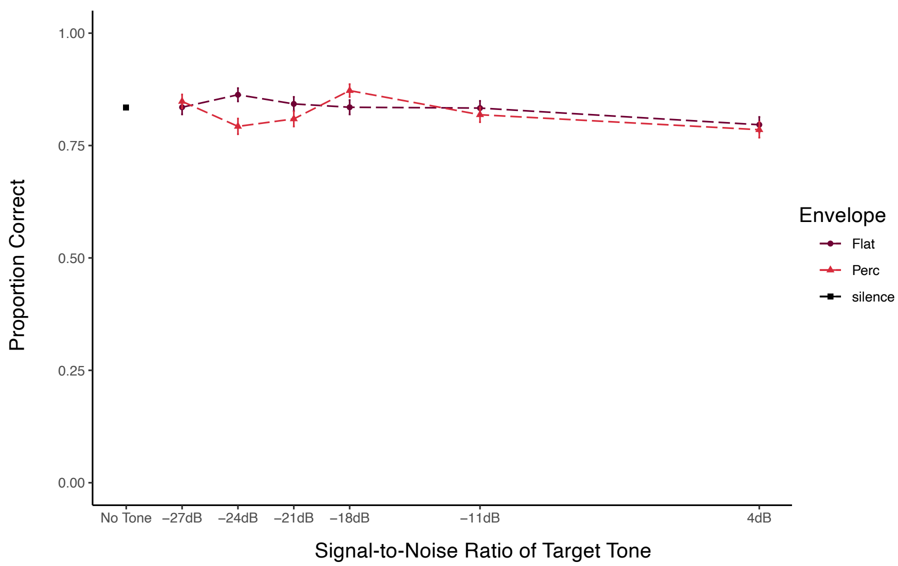
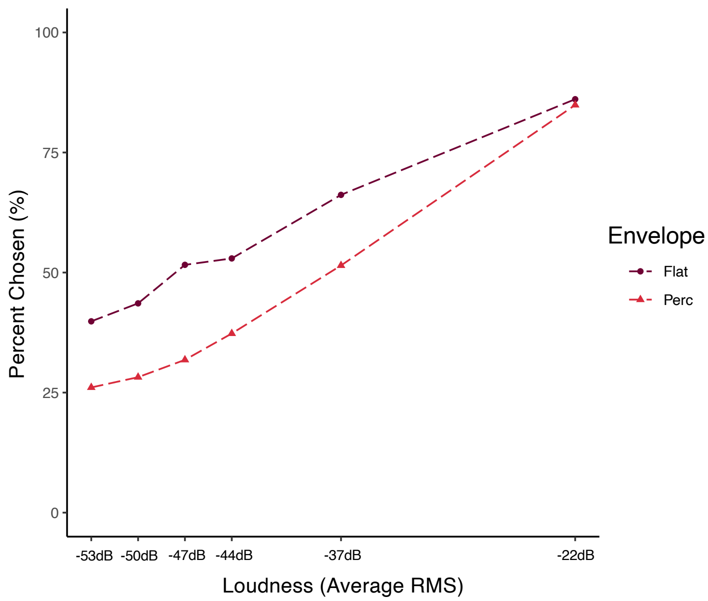


Figure 3. Annoyance Rating Across Envelope and Loudness

Figure 3.



Chapter 4

General Conclusion

Together, these investigations show that temporal variation in acoustic properties can reduce annoyance without hindering detectability. This provides further evidence that temporal variation is both an important factor in understanding our perception of sound (Schutz & Gillard, 2020), and a fruitful design parameter for applications such as auditory alarms (Foley, Anderson, & Schutz, 2020). In chapter two we found that temporal variability in a sounds spectral structure and envelope significantly decreased annoyance while maintaining detection accuracy, and even improving detection accuracy in low levels. The detection advantage observed is consistent with evidence showing temporal fluctuations aid in segregation (Bendixen, Denham, Gyimesi, & Winkler, 2010; Iverson, 1995; Moore & Gockel, 2002) and reduces masking (Buus, 1985). These findings are also consistent with previous investigations showing temporally variable envelopes in auditory alarms can reduce annoyance without harming learnability of alarms (Sreetharan, Schlesinger, & Schutz, 2021).

In chapter three we extended our investigation of envelope by studying tone detectability with a concurrent speech task to better simulate applications such as auditory alarms. Consistent with chapter two we found temporally variable percussive tones to be rated as significantly less annoying while maintaining detection accuracy in relatively high signal levels and improving detection in low signal levels. By employing a dual-task paradigm, chapter three further strengthens our finding that temporal variation would improve the efficacy and usability of auditory alarms.

Further experimentation is needed to clarify the underlying mechanisms of our observed detection advantage in chapters two and three. Separating the role of peak amplitude, and overall variation in envelope will help clarify the difference between percussive and flat tones in low signal level detection. Additionally future directions exploring differences in perceived intensity of stimuli such as those used here will help contextualize these results. Although efforts were made to equate overall loudness by matching average RMS values, further loudness matching experiments are needed to ensure this.

Combined together chapters two and three show that temporal variation in acoustic properties, something present in natural sounds (Foley & Schutz, 2021), may help the usability of auditory alarms. By reducing annoyance without harming the detectability, alarms can be more useful and effective to users. Employing perceptually validated alarm sounds will help create environments that not only sound better, but are safer.

References

Bendixen, A., Denham, S. L., Gyimesi, K., & Winkler, I. (2010). Regular patterns stabilize auditory streams. *The Journal of the Acoustical Society of America*, *128*(6), 3658–3666.

Buus, S. (1985). Release from masking caused by envelope fluctuations. *Journal of the Acoustical Society of America*, *78*(6), 1958–1965. https://doi.org/10.1121/1.392652

Foley, L., Anderson, C., & Schutz, M. (2020). Re-Sounding Alarms: Designing Ergonomic Auditory Interfaces by Embracing Musical Insights. *Healthcare*, *8*(4), 389–389. Retrieved from https://www.mdpi.com/2227-9032/8/4/389

Foley, L., & Schutz, M. (2021). High Time For Temporal Variation: Improving Sonic Interaction With Auditory Interfaces. *IEEE Signal Processing Magazine*.

Iverson, P. (1995). Auditory stream segregation by musical timbre: Effects of static and dynamic acoustic attributes. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(4), 751–763. https://doi.org/http://dx.doi.org/10.1037/0096-1523.21.4.751

Moore, B. C. J., & Gockel, H. E. (2002). Factors Influencing Sequential Stream Segregation, *88*(November 2001), 320–332.

Schutz, M., & Gillard, J. (2020). On the generalization of tones: A detailed exploration of non-speech auditory perception stimuli. *Scientific Reports*, *10*(1), 9520. https://doi.org/10.1038/s41598-020-63132-2

Sreetharan, S., Schlesinger, J. J., & Schutz, M. (2021). Decaying amplitude envelopes reduce alarm annoyance: Exploring new approaches to improving auditory interfaces. *Applied Ergonomics*, *96*.