THE IMPORTANCE OF AMPLITUDE ENVELOPE: SURVEYING THE TEMPORAL STRUCTURE OF SOUNDS IN PERCEPTUAL RESEARCH

Jessica Gillard
McMaster Institute for Music and the Mind
gillarj@mcmaster.ca

Michael Schutz
McMaster Institute for Music and the Mind
schutz@mcmaster.ca

ABSTRACT

Our lab's research has repeatedly documented significant differences in the outcomes of perception experiments using *flat* (i.e. sustained) vs. *percussive* (i.e. decaying) tones [1, 2]. Some of these findings contrast with well-established theories and models, and we suspect this discrepancy stems from a traditional focus on *flat* tones in psychophysical research on auditory perception. To explore this issue, we surveyed 94 articles published in *Attention, Perception & Psychophysics,* classifying the temporal structure (i.e. amplitude envelope) of each sound using five categories: *flat* (i.e. sustained with abruptly ending offsets), *percussive* (i.e. naturally decaying offsets), *click train* (i.e. a series of rapid sound-bursts), *other*, and *not specified* (i.e. insufficient specification with respect to temporal structure). The use of *flat* tones (31%) clearly outnumbered *percussive* (4.5%). This under-utilization of *percussive* sounds is intriguing, given their ecological prevalence outside the lab [3,4]. Interestingly, 55% of the tones encountered fell within the *not specified* category. This is not indicative of general neglect, as these articles frequently specified other details such as spectral envelope, headphone model, and model of computer/synthesizer. This suggests that temporal structure's full importance has not traditionally been recognized, and that it represents a rich area for future research and exploration.

1. INTRODUCTION

Research in the field of audition has a long history of using artificial (i.e. sustained or '*flat*') tones to assess perceptual and cognitive ability. While these *flat* tones lend themselves well for the kinds of rigorously controlled stimuli desirable in an experimental or clinical setting, they offer little resemblance to the types of sounds heard outside the laboratory or audiologist's office [5].

1.1 Stimuli used in auditory perception research

In broad strokes, this issue has been addressed in the literature previously by William Gaver [3, 4], who argued that auditory perception research largely focuses on specific *attributes* of sounds, such as pitch, loudness or timbre. This contrasts with our listening outside the laboratory, which is generally focused more on the *events* producing sounds. Gaver referred to this event-based perception as 'everyday listening' – conveying its pervasive nature in real world settings. For instance, when listening to two hands colliding, one might remark that it 'sounds like hands clapping', not that 'it sounds like a spectrally dense burst of noise with a sharp onset' [6]. Such scenarios can also be observed in laboratory settings – in free identification tasks, participants often describe sounds based on the events creating them rather than their attributes (unless the source is ambiguous) [7].

Such event-based perception can be derived in part from a sound's temporal structure or amplitude envelope. Impact sounds such as handclaps, footsteps and door slams are pervasive in our environment and carry detailed information regarding the materials and force used, particularly in their offset. While this information is easily derived from ecologically valid impact sounds, this is not the case for the abruptly ending *flat* tones commonly used in auditory perception research. In fact, previous studies in our lab have repeatedly shown striking differences in outcomes when using sounds with abruptly-ending *flat* vs. more naturalistic, gradually decaying '*percussive*' tones in a variety of tasks [1, 2]. Examples of *flat* and *percussive* tones used in those experiments can been seen in Figure 1.

1.2 Temporal structure and sensory integration

Our interest in this issue began with a seemingly unrelated debate among percussionist in which some argue that stroke length can influence perceived note duration, with longer gestures making 'long' notes and shorter gestures making 'short' notes. To test this hypothesis empirically we asked participants to rate the durations of tones paired with videos of a professional marimbist making either long or short striking gestures, while ignoring the visual information [1].

When paired with *flat* tones, the visual information (i.e. gestures) did not influence perceived tone durations.

However, when paired with *percussive* tones, long gestures made the tones sound 'longer' and short gestures made the tones sound 'shorter'. Curiously, the use of sounds with naturally decaying offset leads to *qualitatively different outcomes* on a seemingly unrelated sensory integration task.

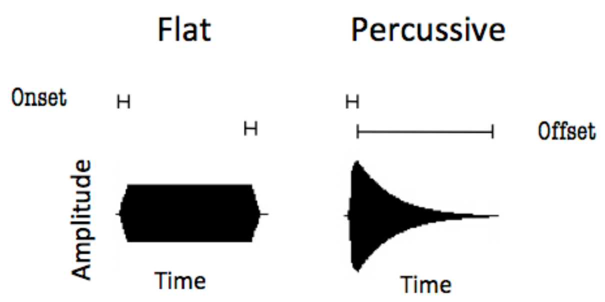

Figure 1. Examples of *flat* (left) and *percussive* (right) tones used in previous experiments [1, 2]. *Flat* tones are characterized by a quick onset, indefinite sustain period and abrupt offset. *Percussive* tones are characterized by a quick onset, followed by an immediate exponentially decaying offset.

This finding is surprising, as it conflicts with the widely held notion that visual information does not affect auditory judgments of duration [8] (although vision is known to affect other aspects, such as localization [9]). This finding has been replicated using point-light displays [10] and a single dot using simplified motion paths [11], suggesting that this discrepancy cannot be fully explained by the use of visual information depicting a marimbist rather than a more traditional visual stimulus. This observation led to our interest in exploring the degree to which this discrepancy can be explained by differences in the perception of sounds with natural vs. artificial envelopes. In other words: is this previously unobserved visual influence on auditory judgments of event duration driven by categorical differences in the perception of ecologically common naturally-decaying sounds vs. the artificial abruptly-ending sounds?

It is worth mentioning that previous studies have used tones with 'ramped' (i.e. increasing in intensity over time) and 'damped' (i.e. decreasing in intensity over time) amplitude envelopes to investigate the perception of streaming vs. bouncing of converging visual stimuli [12]. Overall, damped tones produced the perception of bouncing visual stimuli whereas ramped tones produced the perception of bouncing and streaming equally. This finding is not surprising, as impact events do not typically produce sounds with ramped temporal structures and therefore should not integrate with the visual stimuli. Likewise, if we attempted to replicate this experiment using *flat* and percussive tones, we would expect fewer 'bounce' responses for *flat* tones as *flat* temporal structures are not indicative of an impact event.

1.3 Duration judgment strategies

Intriguingly, differences in the outcomes of perceptual tasks involving *flat* and *percussive* tones are not limited to sensory integration. Other members of our team have found evidence for the use of different strategies when estimating the duration of *flat* vs. *percussive* tones [2]. With *flat* tones participants are able to use what we call a 'marker strategy', marking the onset and offset of a tone to derive the duration. Consistent with the pacemaker-accumulator model [13], participants may be neurally tracking the accumulation of time-markers between the onset and offset of *flat* tones. Such approaches are ill-suited to frequently encountered *percussive* tones, where we suspect participants might use what we refer to as a 'prediction strategy' in which an estimation of the moment of tone completion can be derived by the rate of offset decay.

When we presented these two types of tones uniformly blocked, we found no difference in the precision of duration judgments, suggesting that participants could easily adopt one strategy over the other. However, when we mixed *flat* and *percussive* tones within a block, participants performed significantly worse on duration estimations of *percussive* tones. In other words, when participants are unable to predict what tone type will be presented in the next trial, they cannot select the optimal strategy. Instead, participants presumably resorted to the 'marker strategy' – a viable but less optimal tactic for estimating the duration of decaying *percussive* tones.

These findings of perceptual differences in both audiovisual integration and tone duration estimation tasks raises the question of whether we process sounds with *percussive* temporal structures in a categorically different way than the *flat* tones commonly used in a research setting. Together, this work (along with other differences observed between *percussive* and *flat* tones in an associative memory task [14]) motivated us to explore the temporal structure of sounds used in auditory perception research. As part of a large-scale effort by several members of our research team, here we surveyed the sounds in one prominent journal, in order to determine the relative prevalence of *flat* vs. *percussive* tones.

2. METHOD

We chose to use *Attention, Perception, & Psychophysics* (formerly *Perception & Psychophysics*) as the basis for our survey, with the intention of selecting roughly one hundred articles focused on human perception of non-speech sounds. Searching *PsycInfo* using the terms '*Perception & Psychophysics*' (Publication), 'Auditory' (Identifier/Key Word) and NOT 'Speech', 'Language', 'Phonetic' and 'Dialect' (Identifiers/Key Words) yielded 422 articles. From this pool we composed our sample by selecting the first two articles from each year of the publication (1966-2012), for a total of 94 articles.

Article	Experiment Num	Sound categories	Functional category	Point weighting	Envelope category
Radeau & Bertelson, 1978 [16]	1	1	stimulus	1.0	percussive
Shinn-Cunnigham, 2000 [17]	1	1	target	1.0	click train
Five experiments, each using	2		target	1.0	click train
a single type of sound	3		target	1.0	click train
	4		target	1.0	click train
	5		target	1.0	click train
Boltz, Mashburn, Jones & Johnson, 1985 [18]	1	2	stimulus	0.5	flat
Two experiments, each using	1		warning tone	0.5	not specified
two types of sounds	2	2	stimulus	0.5	flat
	2		warning tone	0.5	not specified
Stilp, Alexander, Keifte & Kluender, 2010 [19]	1	3	target A	0.33	other
Two experiments, each using	1		target B	0.33	other
three types of sounds	1		precursor	0.33	not specified
	2	3	target A	0.33	other
	2		target B	0.33	other
	2		precursor	0.33	not specified

Table 1. Each experiment received a single point, which we distributed equally amongst the functional categories of the sounds used.

We modeled our approach of sound categorization after an earlier survey of articles in the journal *Music Perception* conducted by the MAPLE Lab [15]. As in that study, we classified only the auditory components and coded all experiments (n=212) in the 94 articles individually. We allocated one point to each experiment, subdividing based on the number of sound categories employed. For example, we allocated a point weighting of 1 to experiments using a single sound. If an experiment contained two sounds (i.e. a target and a probe), we allocated each sound category a point weighting of 0.5. If these categories contained multiple sounds, we split the category's weighting equally. Examples of point weight distributions are illustrated in Table 1 [16:19].

2.1 Categories of primary classification

We classified each sound into one of five categories: (1) *flat*, (2) *percussive*, (3) *click train*, (4) *other*, and (5) *not specified*. We classified sounds as *flat* if the description included a period of sustain with rise/fall times. For example, Watson and Clopton had a "550-Hz sinusoid, 150msec in duration…gated with a rise-decay time of 25msec" [20] (suggesting a sustain period of 100msec). Other examples of *flat* descriptions included more ambiguous descriptions such as 'fade-ins and fade-outs to avoid clicks' [21], which imply rise/fall times of an unspecified duration.

We classified sounds as *percussive* if they consisted of a sharp onset followed by a period of exponential decay. Although rarely explicitly described this way, *percussive* temporal structures are implied by the sound produced by certain instruments and/or materials. Therefore, we included studies using traditional percussive sounds such as cowbell [22], chimes [23], bells [23], and bongos [16], as well as impact sounds such as footsteps [24], hand claps [23] and objects dropped on a surface [25] in the *percussive* category. Additionally, we classified piano tones [22] as *percussive* given that they are produced by impact events (i.e. a hammer striking a string).

We classified sounds as *click trains* if they consisted of a series of repeated stimuli over a short duration. In most cases, these stimuli were explicitly described as 'click trains' or 'pulses in a train'. One study described its sounds as "a series of free-field acoustic clicks" [26], which we also included in this category.

The *other* category encompassed sounds with specified envelopes other than those described previously. This included natural sounds such as recordings of complex environmental sounds [23], and tones produced by brass and wind instruments [19, 23], as well as artificial sounds such as amplitude modulated tones [27] and 'pyramid' tones (i.e. with linear rise/fall times but no sustain period) [28, 29].

We treated our final classification of *not specified* as a 'category of last resort', used only when the information regarding temporal structure was insufficient to classify stimuli into one of the previous four categories. For cases in which stimuli were *not specified* in their description but available online, we simply downloaded the stimuli and classified them accordingly. In the current survey, one paper failing to specify the temporal structure included a link to a webpage containing the stimuli. Therefore we determined the envelope shape by analyzing these files [23].

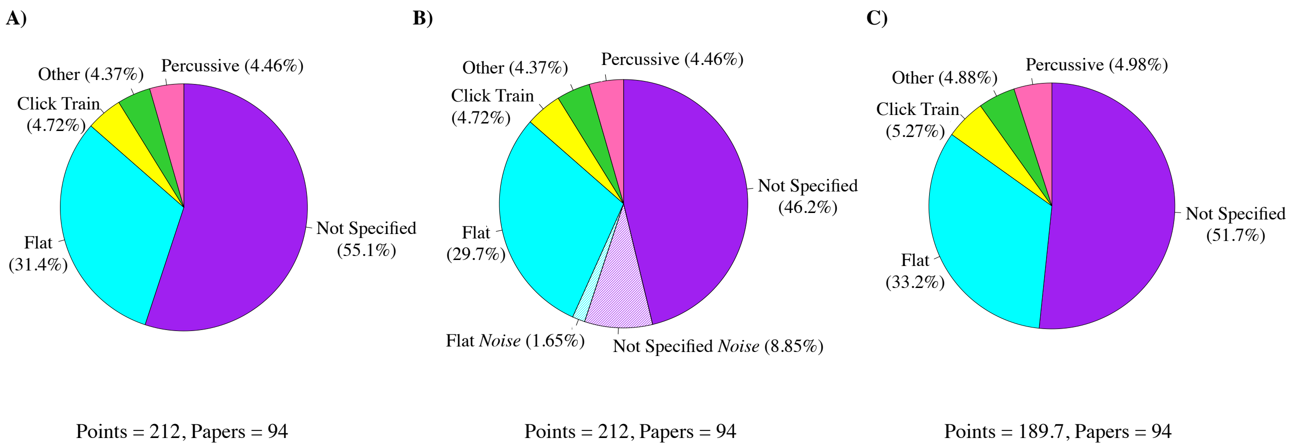

Figure 2. Distribution of the temporal structures of sounds used in *Attention, Perception & Psychophysics*, 1966-2012. Depicts the distribution of envelope types when (A) masking or background noise is included, (B) separated and (C) removed (note the smaller number of "points" included in panel c). Ultimately, decisions regarding the classification of noise do not meaningfully alter the results of this survey.

Amplitude modulated tones presented an interesting classification challenge. At slow rates (i.e. 3Hz) modulators can effectively change the amplitude envelope from *flat* to a sinusoid-like shape. However, at faster rates (i.e. periods sufficiently shorter than the tone's duration) the modulators are too fast to play a significant role on amplitude envelope. Only one paper in our sample used a slow modulation rate (3Hz [27]); the rest exceeded 20Hz. Therefore we placed the 3Hz amplitude modulated tones in *other* category, and classified the remaining amplitude modulated tones (with modulators ranging from 80Hz to 200Hz) according to the previously stated criteria.

2.2 Secondary classifications

In addition to our main focus on temporal structure, we classified other important stimulus characteristics such as the spectral structure, duration, and intensity. Additionally, we noted descriptions of technical equipment information such as headphone model and the model of sound generators so as to gauge methodological diligence. We used the following criteria for this supplemental information:

2.2.1 Spectral information

Spectral information included descriptions such as pure/sine tones, complex tones, white noise, amplitude modulated tones, etc. If an instrument produced a sound, we simply used the instrument name as gross aspects of their spectral structure are already well known. We used the category *not specified (spectral information)* when the spectral structure was not given.

2.2.2 Duration

Duration simply tracked the temporal length of the sound. In the case of *click trains*, we recorded durations in two ways: the duration of individual clicks (if given) and the duration of the train. We denoted descriptions of sounds that did not include duration information as *not specified (duration)*.

2.2.3 Intensity

We recoded the intensities of sounds in decibels, or if described as being presented at a 'comfortable listening level'. If insufficient information was provided regarding sound pressure level, we recorded these sounds as *not specified (intensity)*.

2.2.4 Equipment

As with the spectral information, duration and intensity, we recorded the make and model of equipment (i.e. headphones, speakers, tone generators and computers) given in experimental descriptions. If no information was provided, we recorded these as *not specified (equipment)*.

3. RESULTS

3.1 Fairly accounting for masking noise

We came across several instances where experiments made use of a target or signal in addition to masking or background noise. In these experiments we often found that the target or signal would be specified in terms of its temporal structure, but the masking or background noise would not. Mindful of the possibility of artificially inflating our *not specified* category, we plotted the survey data in three ways: including the masking/background noise (Figure 2A), specifying the proportions of background/masking noise separately (Figure 2B) and, removing it from the sample entirely (Figure 2C). As seen in Figure 2, these considerations did not significantly change the outcome of the survey. Therefore, all reported results are based on the full survey, counting all data points including background and masking noise.

3.2 Outcomes of classifications

As seen in Figure 2A, the majority of sounds (55.1%) used in *Attention, Perception & Psychophysics* fell within the *not specified* category in terms of their temporal structure. This finding echoes the results of an earlier survey classifying the types of sounds used in the journal *Music Perception*, in which *not specified* also formed the largest category [15]. As in that survey, we speculate that many of the *not specified* tones actually have *flat* temporal structures. If so, then tones with artificial abruptly ending offsets comprise 93% of the experiments surveyed (i.e. *not specified*, *flat*, *click trains* as well as amplitude modulated and pyramid tones classified in the *other* category), with less than 7% of experiments making use of sounds with more naturalistic offsets (with *percussive* sounds making up almost two thirds – 4.46%).

This disproportionate focus on tones with abruptly ending artificial envelopes rather than sounds with envelopes more indicative of those encountered in everyday perceiving raises intriguing questions about the generalizability of this research to real world listening.

3.3 Specifications for the *not specified* tones

It is worthwhile to note that the lack of specification of temporal structure is not an indication of a lack of attention to detail on the part of the authors. In fact, within the *not specified* category 61.9% of studies specified another important acoustic parameter known to affect perception – spectral structure. Intriguingly, 66.5% of studies in the *not specified* category denoted the exact model of speaker (i.e. "Sony SRS-A91", "Harman/Kardon HK-195", "Acoustic Profile PSL 0.5") or headphones (i.e. "Sennheiser HD465", "Sony MDR CD250", "AKG-K270", "Beyer DT-49"), used for the experiment and 52% specified the precise model of tone generator (i.e. "Grason-Stadler 455C noise generator", "Hewlett-Packard 200 ABR oscillator", "Wavetek Model 116 oscillator") used to produce the stimuli.

These comparatively high proportions of specification with respect to spectral structure and technical equipment indicate that authors and reviewers frequently felt compelled to include in-depth technical information. Therefore the lack of specification of regarding temporal structure does not indicate negligence to detail, but rather demonstrates that temporal structure has not been previously recognized as playing a meaningful role in the outcome of perceptual experiments—or at least a role less significant than that of the specific model of headphone used to deliver sounds, or the specific tone generator used to synthesize a pure tone.

4. DISCUSSION

Previous work conducted by our research team has repeatedly shown that a sound's temporal structure has the ability to qualitatively change the outcome of perceptual experiments. In an audiovisual integration task, long and short striking gestures did not influence tone duration estimations when paired with *flat* tones. However, when paired with *percussive* tones, short gestures made the tones sound 'shorter' and long gestures made tones sound 'longer' [1]. Similarly, it appears that participants adopt different strategies to estimate the duration of *flat* vs. *percussive* tones [2]. When tones are uniformly blocked, participants adopt a 'marker strategy' – marking the beginning and end – to estimate the duration of *flat* tones and a 'predictor strategy' – deriving an estimation based on the rate of decay – for *percussive* tones.

Beyond perceptual experiments, we have also found differences in performance on a more cognitive task involving associative memory of tone sequences [14]. In this experiment we asked participants to associate ten everyday objects with ten short 4-note melodies that were either *percussive* or *flat*. The results indicate that participants not only learned the associations faster, but could recall significantly more object-associations when hearing *percussive* tone sequences.

These findings suggest that we process sound with naturally decaying offsets in a qualitatively different way than the abruptly ending *flat* tones. While impact sounds exhibiting naturally decaying temporal structures are pervasive in our everyday listening [3], these types of sounds have not historically been used in auditory perception research. Instead, *flat* sounds with artificial sustain periods and abrupt offsets appear to dominate psychophysical research on auditory perception. This survey helps to test our intuitions regarding a long-standing bias towards sounds with artificial envelopes by establishing a set of data capable of commenting on this issue of broad relevance to the auditory perception community.

As anticipated, the proportion of *flat* tones (31.4%) significantly outnumbered the proportion of *percussive* tones (4.46%) in the current survey. This finding differs from our previous survey of the journal *Music Perception*, in which the proportions of *percussive* and *flat* tones were almost equivalent (i.e. 26.9% for *percussive* and 27.6% for *flat*) [15]. This difference may be due to a larger focus on listening to natural (i.e. acoustic) sounds in the case of *Music Perception*, compared to a more psychophysical approach in the case of *Attention, Perception & Psychophysics*. Despite these differences, we did find that *not specified* formed the largest proportion of sounds within each journal. In the current survey, *not specified* in fact encompassed the majority of the sample – 55.1%. This is appreciably more than the 35% encountered in *Music Perception*. Therefore these findings extend our previous results, demonstrating that the general lack of amplitude envelope specification is not limited to a single journal. The degree to which this is a widespread issue within the field of auditory perception remains an open question—albeit one we are currently exploring by surveying other journals.

It is also worthwhile to mention that the exploration of the temporal structure of sounds is not a new idea. In fact, over the last few decades a small group of researchers have been conducting experiments that focus specifically on the perceptual differences produced by varying temporal structures. For instance, these researchers are finding that 'ramped' sounds (i.e. increasing in intensity over time) are consistently perceived as longer than 'damped' sounds (i.e. decreasing in intensity over time) of equal duration, when presented alone [30, 31] or accompanied by a visual stimulus [12, 32]. This suggests that the importance of amplitude envelope in auditory perception research is starting to be recognized and has great potential to flourish.

4.1 Further thoughts on the *not specified* tones

Due to the ease with which *flat* tones can be created and the control afforded, we suspect that many of the *not specified* sounds in the current survey are actually *flat* in nature. Working on this assumption and noting that *click trains* exhibit abruptly ending offsets, we suspect that 93% of sounds surveyed used abruptly ending envelopes (i.e. *flat*, *not specified*, *click trains* as well as amplitude modulated and pyramid tones within the *other* category) that may afford different processing strategies and lead to different experimental outcomes than would be obtained using sounds with more natural envelopes. In contrast, *percussive* sounds with naturally decaying envelopes are used in just 4.46% of experiments surveyed, despite their ubiquity in everyday listening [3].

Although we suspect the *not specified* sounds are in fact *flat,* it would be irresponsible to draw any strong conclusions based on this assumption. As this information was not explicitly provided, researchers attempting to replicate the reported results would not be able to recreate the described stimuli. This lack of specification with regards to temporal structure does not however suggest an inattention to methodological detail. Within the *not specified* category, there is evidence authors went to great lengths to rigorously specify important methodological details. For example, a large proportion (66.5%) of studies in the *not specified* category included the exact model of speakers or headphones used to deliver the stimuli – details that could arguably not significantly influence the overall outcome of the experiment.

Additionally, the spectral structure of the sound as well as other equipment details, such as the model of tone generator and computer, were commonly included in the descriptions of *not specified* stimuli. This attention to detail with respect to other aspects of the methodological and technical details suggests that temporal structure has simply not been recognized as a parameter that could inherently influence results.

4.2 Implications for auditory research and future directions

As a result of our lab's findings in audiovisual integration [1] and tone duration estimation [2] tasks using *flat* and *percussive* tones, we have reason to believe that temporal structure can be an influential parameter in auditory perception research. Given the sophistication of modern sound synthesis tools, we now have the ability to generate sounds with more realistic envelopes while still tightly controlling other parameters. Consequently, researchers are well equipped to assess our perceptual system with sounds paralleling those encountered in the real world. To help others interested in exploring these issues, we are now sharing the software we have developed to generate *flat* and *percussive* tones (Figure 1), which is freely available at www.maplelab.net/software. In the future, we plan to expand our survey to other important auditory perception journals such as the *Journal of the Acoustical Society or America* and *Hearing Research*. Ultimately, we believe temporal structure is a parameter with great potential for fruitful future research, and hope that our survey can help inspire interest in this under-studied aspect of auditory perception.

Acknowledgments

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 (ER 10-07-195) the Canadian Foundation for Innovation (CFI-LOF 30101), and the McMaster University Arts Research Board (ARB) program.

5. REFERENCES

[1] M. Schutz, "Crossmodal integration: The search for unity" Doctoral thesis, University of Virginia, 2009.

[2] M. Schutz, G. Vallet, and D. Shore, "Exploring the role of amplitude envelope in duration estimation: Evidence for two strategies" Presented at Auditory, Perception, Cognition, and Action Meeting, Minneapolis, 2012, pp.7.

[3] W. Gaver, "What in the world do we hear? An ecological approach to auditory event perception" in Ecological Psychology, 1993, pp. 1-29.

[4] W. Gaver, "How do we hear in the world? Explorations in ecological acoustics" in Ecological Psychology, 1993, pp. 285-313.

[5] Acoustical Society of America, American National Standard Specification for Audiometers. American National Standards Institute, 2010.

[6] N. Vanderveer, "Ecological acoustics: Human perception of environmental sounds", Doctoral Dissertation, Georgia Institute of Technology, 1979.

[7] W. Gaver, "Everyday listening and auditory icons" Doctoral Dissertation, University of California, San Diego, 1988.

[8] S. Guttman, L. Gilroy and R. Blake, "Hearing what the eyes see: Auditory encoding of visual temporal sequences" in Psychological Science, 2005, pp. 228-235.

[9] C. Jackson, "Visual Factors in auditory localization" in Quarterly Journal of Experiment Psychology, 1953, pp. 37-41.

[10] M. Schutz and M. Kubovy, "Deconstructing a musical illusion: Point-light representations capture

salient properties of impact motions" in Canadian Acoustics, 2009, pp. 23-28.

[11] J. Armontrout, M. Schutz and M. Kubovy, "Visual determinants of a cross-modal illusion" in Attention, Perception & Psychophysics, 2009, pp. 1618-1627.

[12] M. Grassi and C. Casco, "Audiovisual bounce-inducing effect: Attention alone does not explain why the discs are bouncing" in Journal of Experimental Psychology: Human Perception and Performance, 2009, pp. 235-243.

[13] C. Buhusi and W. Meck, "What makes us tick? Functional and neural mechanisms of interval timing" in Nature Reviews Neuroscience, 2005, pp. 755-765.

[14] M. Schutz, J. Stefanucci, S. Baum and A. Roth "Name that percussive tune: Associative memory and amplitude envelope", Under Review.

[15] M. Schutz and J. Vaisberg, "Surveying the temporal structure of sounds used in Music Perception" in Music Perception, In Press.

[16] M. Radeau and P. Bertelson, "Cognitive factors and adaptation to auditory-visual discordance" in Perception & Psychophysics, 1978, pp. 341-343.

[17] B. Shinn-Cunningham, "Adapting to remapped auditory localization cues: A decision-theory model" in Perception & Psychophysics, 2000, pp. 33-47.

[18] M. Boltz, E. Mashburn, M. Jones and W. Johnson, "Serial-pattern structure and temporal-order recognition" in Perception & Psychophysics, 1985, pp. 209-217.

[19] C. Stilp, J. Alexander, M. Kiefte and K. Kluender, "Auditory color constancy: Calibration to reliable spectral properties across nonspeech context and targets" in Attention, Perception & Psychophysics, 2010, pp. 470-480.

[20] C. Watson and B. Clopton, "Motivated changes of auditory sensitivity in a simple detection task" in Perception & Psychophysics, 1969, pp. 281-287.

[21] P. Bertelson, J. Vroomen, B de Gelder and J. Driver, "The ventriloquist effect does not depend on the direction of deliberate visual attention" in Perception & Psychophysics, 2000, pp. 321-332.

[22] P. Pfordresher and C Plamer, "Effect of hearing the past, present, or future during music performance" in Perception & Psychophysics, 2006, pp. 362-376.

[23] M. Gregg and A. Samuel, "The importance of semantics in auditory representations" in Attention, Perception & Psychophysics, 2009, pp. 607-619.

[24] R. Pastore, J. Flint, J. Gaston and M. Solomon, "Auditory event perception: The source-perception loop for posture in human gait" in Perception & Psychophysics, 2008, pp. 13-29.

[25] M. Grassi, "Do we hear size or sound? Balls dropped on plates" in Perception & Psychophysics, 2005, pp. 274-284.

[26] W. Uttal and P. Smith, "Contralateral and heteromodal interaction effects in somatosensation: Do they exist?" in Perception & Psychophysics, 1967, pp. 363-368.

[27] L. Riecke, A van Opstal and E Formisano, " The auditory continuity illusion: A parametric investigation and filter model" in Perception & Psychophysics, 2008, pp. 1-12.

[28] B. Wright and M. Fitzgerald, "The time course of attention in a simple auditory detection task" in Perception & Psychophysics, 2004, pp. 508-516.

[29] E. Hasuo, Y. Nakajima, S, Osawa and H. Fujishima, "Effect of temporal shapes of sound markers on the perception of interonset intervals" in Attention, Perception & Psychophysics, 2012, pp. 430-445.

[30] R. Schlauch, D. Ries and J. Di Giovanni, "Duration Discrimination and subjective duration for ramped and damped sounds" in Journal of the Acoustical Society of America, 2001, pp. 2880-2887.

[31] M. Grassi and C. Darwin, "The subjective duration of ramped and damped sounds" in Perception & Psychophysics, 2006, pp. 1382-1392.

[32] J. Neuhoff, "Perceptual bias for rising tones" in Nature, 1998, pp. 123-124.