Auditory Discrimination of Pure and Complex Waveforms Combined With Vibrotactile Feedback

Gareth W. Young  
Dept. Computer Science / Music  
University College Cork  
Cork, Ireland  
g.young@cs.ucc.ie

David Murphy  
 Dept. Computer Science  
University College Cork  
Cork, Ireland  
d.murphy@cs.ucc.ie

Jeffrey Weeter  
 Dept. Music  
University College Cork  
Cork, Ireland  
j.weeter@ucc.ie

ABSTRACT

Here we present an experiment that investigates the application of extra-sensory vibrotactile stimulus of pure and complex waveforms in audio frequency difference detection. Vibrotactile information was expected to have some influence upon the frequency discrimination of audio Just Noticeable Difference (JND) tests. The experiment explored the effects of vibrotactile feedback in combination with audio frequency discrimination for the same f0 of 160 Hz. The potential of a correctly identified result for two separate groups were measured and then compared. Group A subjects were given an audio only JND test via headphones, whilst group B were given the same test with additional vibrotactile stimulus delivered via the Audio-tactile Glove. The results of our experiments illustrate that vibrotactile feedback can affect our ability to perceive changes in pitch when presented asynchronously at f0.

Author Keywords


ACM Classification

H.5.2 [Information Interfaces and Presentation] User Interfaces, Haptic I/O, Auditory (non-speech) feedback H.5.5 [Information Interfaces and Presentation] Sound and Music Computing.

1. INTRODUCTION

The manner in which auditory and haptic cues are integrated into musical performances are detailed in the findings of a number of recent studies, outlining the role therein of human senses beyond that of the auditory modality [1, 2, 3, 4]. Research has also shown that the neural substrates of both the auditory and tactile systems are shared at a much lower level than previously understood [5, 6]. For our experiments, we have chosen to focus on the cumulative effects of both vibrotactile feedback applied in tandem with an auditory stimulus. Audio and tactile feedback occurs in most instances of human interaction that involve touch, but it rarely occurs at a cognitive decision making level. This cross-modal effect has been demonstrated in the tactile illusions that transpire from the modification of related audio stimuli, as seen in the “Parchment-skin illusion” [7]. It can therefore be observed that an audio stimulus can be modified to alter a tactile experience.

The function of extra-auditory cues and their input into the field of perceptual materials has been a major contributory part of investigations into how music is perceived. These include the influences of tactile and auditory feedback upon a performer [8], the performer’s understanding of the musical structure of a piece of music [9], and the portrayal of a score’s content [10]. The conclusions found in such research have suggested that multimodal sensory cues are responsible for indirectly augmenting the auditory perception of music.

2. Background

Recent studies have shown evidence of interaction between the auditory and the somatosensory systems at a multitude of stages within the human central nervous system [11]. The combination of these two sensory modalities exceeds the predicted unisensory summation of the two stimuli alone, proving that multisensory convergences occur at a much lower level than previously thought. We have observed enhancements in auditory processing through the addition of tactile feedback, which elevates the response speeds to those of suprathreshold stimuli [12]. We have also observed improvements in the intensity perception of faint tones [13]. These finding contradict earlier research, which stated that auditory detection is impaired by simultaneous tactile stimuli and vice-versa [14, 15, 16]. Early research indicated that tactile stimulus could be seen to enhance simultaneous auditory stimuli, but the interpretation of results pointed towards an affection of response criteria and perceptual sensitivity due to increases to both signal and noise [17]. Other studies indicate that the detection of a stimulus is enhanced when it simultaneously registers with two or more sensory modalities [18]. Reinforcement of neural activity occurs when two modalities stimulate a person in near unison of time and placement. Unlike visual data observations, haptic cues in a musical performance are captured via contact with vibrating sound-emanating objects.

During a music performance, the control mechanisms of the performer rely on feedback produced by an acoustic instrument [19]. This feedback presents itself to the internal senses of the musician so that they are able to adjust and maneuver their bodies in response. The transmission of vibration to the performer is an integral feature that directly relates to the design requirements of acoustic musical instruments. Regardless of the manner of human affecter placement, the feedback remains almost constant [19, 20]. In music, auditory and tactile communications result from sensory stimulation via physical mechanical pressure in the form of oscillations [20, 21]. The interactions of mechanical vibrations against the mechanoreceptors of the skin and within the cochlea activate neural impulses to be processed by the brain. The closeness in relationship between the neural processing of these two modalities of transduction has been proven in aforementioned research [21]. Our examination here focuses on the combined feedback received by the auditory and tactile systems.

We have chosen to focus our current study on audio frequency tactile stimulus as a supporting sensory input. Both audio and tactile stimuli overlap in the same frequency range. However, the limitation of interaction involving both hearing and touch are restricted by the increased sensitivity of the ear in
comparison to the skin. On average, the ear can sense a range of approximately 20 to 20 kHz, while skin can only sense a narrower range of 0.3 to 1 kHz. Tactile and auditory information, whether applied concurrently or in an alternating presentation, may be perceived as interleaved signals [22]. Within this range, vibrotactile information has been shown to stimulate the auditory cortex [23].

Experiments with the Audio-Tactile Glove have presented results in tactile detection levels, including complex waveforms [24]. These values were used to minimize perceived amplitude differences in waveforms for our current experiment. As can be seen in Figure 1, the sub-threshold of vibrotactile stimulus detection can be divided into distinct sub-ranges, pertaining to the frequencies that are cutaneously detectable and the waveform being applied. The stimuli presented during experiment at $f_0$ were delivered with the adjusted output amplitudes dependent on the waveform; they were also applied in synchronous phase. Previous research has shown that the auditory and vibrotactile systems combine whilst performing objective detection tasks, regardless of the relative phase and non-temporal synchrony [25]. These results indicate that both of the neural pathways of the auditory and tactile systems combine through a common or related network. Therefore, we kept relative phase constant by delivering stimuli from the same source.

3. Pitch Discrimination of Pure and Complex Waveforms

Our experiment sought to investigate the relationship between vibrotactile feedback (at an audible frequency) and a subject’s ability to distinguish audible changes in frequency at the fundamental. Due to the sensory dominance of hearing over tactile, we requested our participants to focus on the audio stimulus only.

![Figure 1: Threshold of perception of vibration applied via the audio-tactile glove](image)

3.1 Stimuli

The audio and vibrotactile stimuli delivered during all experiment conditions were sine, saw and square waveforms with a fundamental frequency ($f_0$) of 160 Hz ($S_1$, $S_2$, and $S_3$ respectively). This was followed by a variation of ± 12 Hz around this fundamental. Subjects were divided into two groups. For group A, dual mono audio only stimuli were delivered via Sennheiser HD 215 headphones. For group B, dual-mono audio and vibrotactile waveforms were delivered to both the hands and ears of each subject via the Audio-Tactile Glove and the same Sennheiser headphones. The signal was applied to both hands simultaneously in order to control for increased dominant hand sensitivity and other variances of hand sensitivity that may have existed. Waveforms were outputted via a digital-analogue sound card (Avid Fast Track C400) with a sampling frequency of 96 kHz and 24-bit resolution. The audio output was routed to the left and right glove in parallel. Waveform clips were digitally generated (Audacity) and recorded. Waveforms were randomly selected from an audio library for replay. Samples were arranged into five-second clips. Each clip consisted of a 2-second waveform sample at $f_0$, a one second inter-stimulus time (IST), and then another 2-second waveform sample that varied in frequency from the first by ± 0.25, 0.5, 0.8, 1, 1.5, 2, 3, 4, 6, 8, 12 Hz.

3.2 Subjects

All participants were randomly divided into two groups (A/B). Participants were then asked to identify themselves as musicians or non-musicians based upon having been formally trained or actively performing regularly in the last five years. Group A consisted of 10 males and 5 females; all aged 22 to 49 ($MD = 28$; $SD = 8.79$). In Group A, 7 participants self identified as musicians and 8 as non-musicians. Group B consisted of 8 males and 7 females; all aged 21 to 40 ($MD = 28$; $SD = 6.26$). In Group B, 10 participants self identified as musicians and 5 as non-musicians. Physiological pre-testing was not performed on individual subjects; however, participants self reported as having no hearing difficulties or other impairments.

3.3 Experimental Conditions

This experiment was designed to measure the pitch perception abilities of the subject for both pure and complex waveforms with a fundamental frequency ($f_0 = 160$ Hz) and highlighted the individual’s Just Noticeable Difference (JND) audio range at this frequency. The waveform pairs were presented in counterbalanced order. Subjects self-evaluated the required loudness for headphone volume and for Group B the same signal was then adjusted to fit within the sub-threshold tactile range for the appropriate waveform. All test subjects were asked to evaluate the relative pitch of two short audio samples. Group A received audio only stimulus via headphones, whilst Group B received both audio and tactile stimulus. Group B were asked to concentrate on the audio stimuli only. The two tones were never presented to the test subjects at the same pitch. Subjects were required to judge if the second clip was higher or lower in pitch than the first by stating “higher” or “lower”. Participants were not permitted to make “same” judgments.

3.4 Results

A Wilcoxon Signed Rank Test revealed that there was no statistically significant effect in the presentation of same frequency variations in positive and negative directions around $f_0$ ($p > 0.05$ for all frequency pairs). Therefore, it was possible to collapse the results across the complementary pairs. Table 1 shows the probability of correct answers for stimulus pairs after frequency collapsing and bias correction for each waveform shape.

An independent-samples t-test was conducted to compare the probability of correct answers found in Groups A and B within the three different waveform parameters (Table 2). These were deemed to be statistically significant differences in probability scores between the groups over the three waveforms tested. These results suggest that the combined audio-tactile stimulation received in Group B had a positive effect upon a subject’s ability to discriminate between changes in frequency. The magnitude of the differences in the mean score was also deemed to have a significantly large effect.

For $S_1$, $S_2$, and $S_3$ the interaction effect between groupings and musical ability was not found to be statistically significant, for $F (1, 26); S_1 = 0.2, p = .66; S_2 = 0.95, p = .34$, and $S_3 = 0.11, p =$
This indicated that there was no significant effect between musicianship and grouping. There was found to be a statistically significant main effect between the groups, for F (1, 26); S1 = 21.6, p = ns; S2 = 12.01, p = .002; S3 = 5.77, p = .024. However, the effect sizes varied for each waveform; for S1 it was large (partial eta squared = 0.45); for S2 the effect size was large (partial eta squared = 0.32); and for S3 the effect size was smaller (partial eta squared = 0.18). This implies that although musicality did not alter participant probability scores in terms of grouping (A or B), there were significant differences in scores between the two groups.

The PSE is the frequency of which “Higher” response. As can be seen in figures 4, there were notable differences between the two groups across the three waveforms. An independent-samples t-test was conducted to compare the PSE scores for Group A and Group B across the three waveforms. There was found to be a significant difference in S1; PSE Group A (M = 160.67, SD = 1.33) and Group B (M = 159.8, SD = 1.33; t(28) = 1.78, p = .08, two-tailed). The magnitude of the differences in the means (mean difference = 0.86, CI: -0.13 to 1.86) was moderate (eta squared = 0.05). For S3, Group A (M = 160.33, SD = 1.2) and Group B (M = 160.16, SD = 1.045; t (17.89) = 0.52, p = .6, two-tailed). The magnitude of the differences in the means (mean difference = 0.17, CI: -0.52 to 0.87) was small (eta squared = 0.005).

The effect of vibrotactile feedback reaches statistical significance; the actual differences in the mean values were variable in size for each waveform. Therefore, the difference between musicality within the groups appears to be of some practical significance. Additionally, for S1 and S2 waveforms there was a statistically significant main effect for subject musicality; for F (1, 26); S1 = 5.69, p = .03; S2 = 7.92, p = .01; and for S3 =1.19, p = .29 results were not statistically significant. The effect size for all waveforms was small to medium (partial eta squared; S1 = 0.18; S2 = 0.19; S3 = 0.04). This means that the musicians performed differently for each waveform as it was presented. This can be seen in figure 3.

A psychometric analysis of results identified the point of subjective equality (PSE) for each stimulus type (S1, S2, and S3) for each group. The PSE is the frequency of which “Higher” responses were deemed to be of equal prominence as “Lower” and subjects were effectively guessing in order to make an accepted response. As can be seen in figures 4, there were notable differences between the two groups across the three waveforms. An independent-samples t-test was conducted to compare the PSE scores for Group A and Group B across the three waveforms. There was found to be a significant difference in S1; PSE Group A (M = 161.63, SD = 1.25) and Group B (M = 160.88, SD = 0.81; t (28) = 1.97, p = .05, two-tailed). The magnitude of the differences in the means (mean difference = 0.76, CI: -0.03 to 1.6) was moderate (eta squared = 0.07).

However, although there was a difference in PSE frequencies for S1 and S3, there was found to be no significant difference in PSE frequencies for S2 and S3. For S2, Group A (M = 159.8, SD = 1.33) and Group B (M = 159.8, SD = 1.33; t(28) = 1.78, p = .08, two-tailed). The magnitude of the differences in the means (mean difference = 0.86, CI: -0.13 to 1.86) was moderate (eta squared = 0.05). For S3, Group A (M = 160.33, SD = 1.2) and Group B (M = 160.16, SD = 1.045; t (17.89) = 0.52, p = .6, two-tailed). The magnitude of the differences in the means (mean difference = 0.17, CI: -0.52 to 0.87) was small (eta squared = 0.005).

4. Discussion

The results from the experiment show how the detection of a frequency change of ±12 Hz at a fundamental of 160 Hz can be facilitated when a simultaneous cross-modal presentation of stimuli occurs in unison on each hand. Additionally, when tactile feedback is combined with audio stimulation we see a marked improvement in our group’s ability to discriminate between audio frequency pitch variations (Group B) above that of levels expected when audio is presented in isolation (Group A) at 160 Hz. Our findings also highlight an increased ability in frequency discriminate by musicians. These findings are congruent with recent studies that suggest that there is a close relationship between auditory and somatosensory stimulation in the auditory cortex of the brain. This relationship can be seen in fMRI observations that have captured the mapping of audio-tactile co-activation in the auditory belt areas of the brain [6]. Specifically, we have shown increases in the probability of correct identification of audio frequency discrimination by musicians whilst tactile stimuli is applied in relative synchrony of phase and amplitude, as would be experienced in real-world musical audio-tactile interactions with acoustic instruments.

![Figure 2: Boxplots representing probability of correct response.](image)

![Figure 4: Probability of correct answer between groups and participants with musical ability.](image)
These results present interesting observation of changes in expected values of JND. The JND of the tactile system is much broader than that of the audible stimuli within the frequency range covered in our experiment. For example, the expected tactile only JND [%] of a 150 Hz sinusoidal stimulus with amplitude held constant is 18% of the fundamental [26]. Which would equate to 28.8 Hz at the \( f_0 \) of our experiment. Additionally, in an audio only JND experiment we would expect to see a 3 Hz variation in JND for sine-wave frequencies below 500 Hz, and 1 Hz for complex tones [27]. As can be seen in Table 3, the JND results for Group A’s audio only test presented a 3.68% JND for sine waveforms and 3.3% for complex waveforms. In the combined audio-tactile group, we observed a small improvement in JND [%], but no significant differences were found. For S_2, the JND was measured at 2.88% of \( f_0 \) and for S_1 and S_3, we observed a JND of 3.1% and 2.88% respectively.

### Table 3: Just Noticeable Difference [%] for each waveform

<table>
<thead>
<tr>
<th></th>
<th>Group A</th>
<th></th>
<th>Group B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JND [25%]</td>
<td>JND [75%]</td>
<td>JND [25%]</td>
<td>JND [75%]</td>
</tr>
<tr>
<td>S_1</td>
<td>1.58</td>
<td>5.76</td>
<td>3.68</td>
<td>1.39</td>
</tr>
<tr>
<td>S_2</td>
<td>2.2</td>
<td>4.37</td>
<td>3.3</td>
<td>3.52</td>
</tr>
<tr>
<td>S_3</td>
<td>2.94</td>
<td>3.66</td>
<td>3.3</td>
<td>2.82</td>
</tr>
</tbody>
</table>

Our findings support the theory that the simultaneous combination of tactile and audio stimulation positively influences the perceptual frequency discrimination of our sensory system. We attribute this to the low-level integration of these two modalities in the cortical system. The relationships between the strength of the two modes of stimulus should directly relate to psychophysical models based upon human sensory parameters. Numerous examples of singular sensory modality interactions can be cited, but it is rarely the case in musical applications that one singular sense is operating alone in environmental interactions. Many natural events and occurrences seek to compete for our combined sensory attention and a multitude of which are capable of stimulating us in several ways at once. Synchronies in audio-tactile events are particularly ingrained in acoustic musical instrument performances where these combined perceptual aspects are innately integrated.

Future developments in our experimental design will make it possible to analyze the role of exploratory active feedback in operations that include vibrotactile information such as secondary feedback paths. However, it is worth noting that vibration perception in musical interactions makes use of both active and passive feedback during instrument performances. Therefore, the role of an external vibratory force operating outside of the body is arguably as important as the perception of vibration through exploratory actions. Active haptic perception is usually the main focus when designing interfaces, but it should perhaps be complemented with some mode of relatable passive feedback in order to create a more complete multimodal interaction. Further, in musical performance, the application of audio related tactile feedback bridges passive with active feedback as the instrumentalist in motion is actively interacting with the source of the passive feedback. In new musical devices, it is in the decoupling of the gestural controller from sound production module that the role of passive and active feedback becomes separated further. 

Figure 3: PSE for each group and stimulus type

Acknowledgements: The authors gratefully acknowledge the sponsorship of the National Institutes of Health. Research supported by Grant No. R01 DC012295-01A1.

**5. Conclusions**

Here we have investigated the role of vibrotactile feedback and its contribution to the detection of auditory perception of frequency changes at 160 Hz. We have shown from our experiments that vibrotactile feedback can affect our ability to perceive positive and negative changes in pitch when presented asynchronously at a frequency of 160 Hz. We discussed the sensitivity ranges of both systems, highlighting the overlap that occurs between them. In light of this overlap, we discussed research that indicates a relationship between vibration perception and auditory processing in the brain. We tested the JND abilities of two groups of subjects. Group A was given an audio only test, whilst Group B was given the same test with concurrent tactile stimuli that was directly related to the audio stimuli. We found that the group with simultaneous multimodal stimulus was able to correctly identify changes in frequency better than the audio only group. Group B identified 91% of frequency changes successfully, whilst Group A correctly identified only 79%. A psychometric analysis of our experiment results identified the point of subjective equality.
transparent and intuitive to operate during a musical research, to combine multisensory interfaces that are carr
somatosensory integration in frequency discrimination tasks. A psychophysical model is required to reinforce the role of
This will remove any doubt of intra-subject differences. In the final section, we discussed the potential meaning of our findings and their application in relation to musical interactions and DMI design. We maintain that the adoption of a combined psychophysical model is required to reinforce the role of somatosensory integration in frequency discrimination tasks that are carried out on digital devices. This will hopefully allow researchers to combine multisensory interfaces that are transparent and intuitive to operate during a musical performance.

6. REFERENCES