ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Musical Acoustics
Session 2pMU: Musical Preference, Perception, and Processing

2pMU12. Refining the stereo technique for augmented ambience gradient: Improvements in stereo image, spatial envelopment, and mixing flexibility

David J. Tagg* and Kevin Fallis

*Corresponding author's address: Music - Sound Recording, McGill University, 555 Sherbrooke St. W., Montreal, H3A 1E3, QC, Canada, jamietagg@gmail.com

While working on location, recording engineers are often challenged by insufficient monitoring capabilities. Poor (temporary control room) acoustics and/or mandatory headphone monitoring can make judgments regarding microphone choice and placement difficult. This compromised monitoring often leads to timbral, phase, and stereo image problems. We are often forced to compromise between the improved spatial imaging of near-coincident techniques and the attractive acoustic envelopment of spaced omnidirectional mics. This research reviews a new technique: Stereo Technique for Augmented Ambience Gradient (STAAG), which aims to improve stereo imaging, ambient envelopment, and flexibility in the mix. Building on a preliminary study, this research realizes ideal microphone angle/spacing combinations to promote spatial accuracy. In addition, it investigates the quality of the ambient envelopment compared to omnidirectional-based techniques, and the ability of STAAG to allow an engineer to manipulate the direct to reverberant energy ratio during post production without corrupting the stability of the stereo image.

Published by the Acoustical Society of America through the American Institute of Physics
1. INTRODUCTION

When recording sound on location in stereo, the recordist is challenged with capturing the source (or ensemble) in conjunction with the ambient sound of the acoustic space. He or she may decide to include more or less ambience depending on the quality and quantity of the early/reverberant reflections in the space, as well as what is appropriate to the period or style of music. It can be challenging for the recording engineer, often having to work exclusively on headphones, to make decisions about how much or how little of the ambience of the room to include in the recording, thus dictating the microphone position. Headphone monitoring creates a highly detailed and unnatural listening experience, often motivating overuse of reverb, low-frequency problems, and unnatural stereo imaging. Monitoring with speakers in a space not designed for critical listening (often a classroom, green room, or storage space) presents similar as well as additional challenges not addressed in this text.

When recording in highly ambient spaces, the recordist often considers coincident or semi-coincident techniques. While the directional characteristics of the microphones used in these techniques allow better isolation of the source(s), we often lose a sense of acoustic envelopment.¹ Too little reverb and one might choose an AB-Omni or Decca Tree configuration to help emphasize reflected energy. However, in doing so, one sacrifices clarity in stereo imaging. When trying to add ambience to techniques using directional mics, the solution is typically handled in either or both of two ways. The first option is to include a separate pair of microphones further away and in the reverberant field, including ambience in an enhancing, uncorrelated way. However, the discrepancy in time of arrival is not true to any particular location, and can cause phase, timbral color, and localization problems. The second option involves the use of artificial reverb added to the mix via signal processing. While this offers a better alternative to phase and timbral coloration problems, since the acoustic model of the processor has little to do with the original space, matching the timbre and time characteristics of the hall can be difficult. This can result in transient distortion and stereo image ambiguity. In the case of a convolution reverb employing an impulse of the original hall, the issues discussed above (involving an additional microphone array) can again become apparent, as well as a host of other challenges, if such practices are even allowed in a particular space.

Stemming from an initial study of the Stereo Technique for Augmented Ambiance Gradient (STAAG), this paper investigates improved stability in stereo imaging and listener envelopment.² Here, perceived sound source localization is examined within a narrow range of forward-facing microphone angles. It is also ideal, while changes in ambience are made, to have the stereo image remain stable or minimally changed. This research shows that this can be accomplished. This research also evaluates listener envelopment (LEV) for several different angles and microphone types, demonstrating how STAAG approaches the experience of spaced omni techniques in this regard.

1.1 Background

Several stereo mic techniques are available for ambient sound/music recording. While all of these techniques successfully address technical/aesthetic needs, each technique offers certain advantages and limitations. Without discussing the specifics of established techniques which are readily available in the literature, a few traits of the most common are worth mention as they directly relate to the STAAG technique.³ In highly ambient spaces where control and isolation from the abundant acoustic reflections is necessary, one might chose the ORTF technique as an appropriate solution. ORTF uses unidirectional cardioid mics with a 17cm spacing and 110° included angle in an effort to mimic the acoustic properties of the head and the way humans receive sound. This configuration, depending on the off-axis characteristics of the particular microphone used, does a satisfactory job of preserving the cues that are necessary for realistic auditory localization. However, in less reverberant acoustic environments where more reflected sound may be desired, the engineer may be motivated to use mics with less directional selectivity. In these circumstances, a technique such as a Decca Tree or AB with omnidirectional mics may be preferred. While these techniques do include more reflected sound energy, and thus a greater sense of LEV, the lack of directional discretion causes phase characteristics that obfuscate the stereo image. The STAAG technique offers an alternative to these techniques, exploiting the best characteristics of many common configurations. This gives control over the direct to reverberant sound ratio, the ability to adjust this ratio without stereo image distortion, and a convincing sense of LEV.

With regard to human auditory perception, we localize sound in azimuth by way of two main schemes. For high frequencies above roughly 1200Hz, the interaural intensity or level difference (ILD) is used to discern the directional origin of a sound. Since the head casts an acoustic shadow for high frequencies, the intensity of the sound arriving at the ear further away will be attenuated. The varying magnitude of these frequencies gives information
about the location of sound sources. At low frequencies from roughly 1200Hz to 350Hz, we use the phase information available due to the interaural time of arrival difference (ITD) at each ear. Since low frequencies refract around the head, no acoustic shadow is cast, and thus intensity differences provide no useful cue. These wavelengths are significantly long compared to the size of the head and are matched in amplitude. The phase difference between the left and right ears allows us to perceive which side of the head is closer. ILD and ITD cues are exploited in various ways with standard stereo microphone techniques.

The STAAG technique implements four directional microphones. This is configured with forward-facing and rearward-facing pairs of mics with a spacing of 17cm, as seen in figure 1.1. In all cases, the forward-facing mics are first-order cardioid polar pattern. The rearward-facing mics can be varied in directionality, but should be no less directional than first-order cardioid in order to maintain integrity in the stereo image for direct sources located in front. Decorrelation is an important factor in our perception of LEV. To emphasize this, testing was done with cardioid, hyper-cardioid, and shotgun microphones.

In previous research on STAAG, a coarse study investigated a wide range of angular combinations for both the front and rear-facing microphone pairs. This revealed two main results. First, it became clear that narrower configurations of forward-facing microphones presented a more accurate and consistent stereo image of the sound sources. Secondly, it showed that narrow angles can be used successfully with this near-coincident technique, and that recordings done with a 17cm spacing outperformed those with a 30cm spacing. This study investigates a refined range of front angles to confirm the results of the previous study, and offer a platform for improvements to be made in other characteristics that benefit the listener.

**FIGURE 1.1 - Top-down view of STAAG configuration**

In addition to the accuracy of the stereo image, the sense of acoustic envelopment is vastly important in acoustic recording. Many studies have been conducted on this topic with regards to perceptions of concert hall acoustics and speaker playback systems. Since the traditional definitions for envelopment generally apply to interaction between a sound source and an environment, there is no established terminology that applies to headphone listening. An additional intention of this research focuses on headphone listening. The prevalence of iPods, smartphones, and other portable media players leave the majority of music appreciators listening to music that is presented over headphones. Therefore, it seems appropriate that music produced in this era be created not only with speaker playback in mind, but also exploit the advantages of direct-ear presentation. Thus, it follows that if we are to evaluate ambient recording techniques, the sense of acoustic envelopment should be investigated.

Many conflicting definitions for envelopment exist. In order to suit the needs of the headphone evaluations done in this study of STAAG, the authors used Jan Berg’s discussion of these conflicting definitions, and adapted several points. Here, we use the term Isolated Listener Envelopment (ILEV), as defined by the listener's sense of being enveloped or surrounded by sound predominantly produced by late-arriving reflected energy. It should be noted that we will use the more traditional term "LEV" when referring to research focusing on concert hall acoustics and multi-channel loudspeaker systems. This late arriving reflected energy can come from many directions, and the direction from which it arrives influences the strength of this effect. For our purposes, we attempt to a) give the engineer control of the ambiance, and b) present the listener with sound cues that can be reliably presented in azimuth. It is fortunate that the most robust cue for creating LEV comes from lateral reflections. In fact, lateral reflected energy has the most significant influence on our perception of LEV, contributing to more than 50% of the effect when accounting for rear, overhead, and frontal reflections.
1.2 Hypothesis

Semi-coincident techniques offer improved spatial cues for sound source localization. Angles narrower than those found in traditional semi-coincident techniques can present an accurate stereo image, and we investigated an optimal configuration for forward-facing mics. However, directional mics lack the sense of LEV/ILEV that make spaced omni mics a desirable choice. Coincident rearward-facing mics offer the engineer control over the quantity of reverb in the mix that remains accurate to an ideal listening point in the hall. Through the use of various angles and polar patterns, they also offer control and emphasis of ILEV by varying the degree of lateral reflected energy. Earlier studies have addressed the use of cardioid pattern mics for the rear-facing channels of STAAG. While the previous study did not evaluate LEV or ILEV, anecdotal listener reports indicated that ILEV experienced by STAAG was lacking, when compared with theAB-Omni technique. By making use of hyper-cardioid and shotgun mic capsules, late reflected energy in the rear channels can be further decorrelated and emphasized in the mix, presenting an enhanced sense of ILEV to the listener.

2. METHODS

2.1 Overview

Several different microphone configurations were used to record samples of an acoustic ensemble. These samples contained musical elements designed to allow for evaluation of dynamics, acoustic envelopment, stereo width, and sound source localization. In order to evaluate sound source localization, it was necessary to record a short musical excerpt with specific characteristics. The ensemble needed a width that would occupy the majority of the angle of interest (> +/- 40°), and few enough instruments that each location could be quickly localized and evaluated by the listener. Under these constraints, the first movement of Beethoven's String Quartet in F Major (op. 18, no. 1) was selected, then shortened to 55 seconds of music (46 measures) containing only the critical musical elements pertinent to this study. These included dynamic range, exposed features (solos) for every instrument, and homogeneity of the ensemble. Since many combinations of mic position were used, it was unreasonable to ask musicians to play the piece repeatedly for the capture, or to assume that these performances would be consistent from one take to the next. For consistency, a "virtual quartet" was created by recording each instrument separately in isolation, then "performed" by loudspeakers placed on stage from which the instrument would typically perform.

These acoustic instruments produce a different frequency response depending on the direction of sound emission. Additionally, loudspeakers are far more directional than the instruments. Capturing the front as well as the rear of each instrument and presenting these signals with forward and rearward-facing loudspeakers created a more realistic acoustic representation of instruments performing in the hall. Each instrument was captured in a traditional recording studio isolation booth using Schoeps CMC6-xtLin microphones with MK5 (cardioid) capsules. Microphones were positioned toward the front and rear of the instrument at a distance of 25 inches, equidistant from the body at the bridge location. The iso-booth was made as acoustically absorbent as possible, with an estimated RT60 < 60ms. These discrete signals were then reproduced with eight loudspeakers at calibrated levels, presenting the music to the various STAAG configurations, in a typical concert hall used for chamber music performance and recording.

2.2 Ensemble Presentation and Performance Recording

All configurations were recorded in Redpath Concert Hall at McGill University. A series of loudspeakers played back the 8-channel ensemble sample. Three Genelec 8050A (violin I, II, and viola) and one JBL LSR6328 (cello) loudspeakers presented the front of each instrument, while four JBL LSR6328 loudspeakers presented the rear. Supported by a chair, forward-facing speakers sat at an upward angle of roughly 20°. The rear facing speaker placed on the floor below the chairs projected toward the back of the hall. Microphones were placed nine feet in front and four and a half feet above the center of the ensemble at floor level. A semi-circle with a seven foot radius and 30° increments spiked on the floor allowed for repeatable click presentation locations. All speakers were calibrated to perform at an average level of 76dB SPL (90dB SPL peak). After initial calibration, levels and equalization were adjusted subjectively to create a convincing sound for each instrument/speaker and a musical balance for the
ensemble. This was done in conference with the two authors and one masters-level recording student. Once adjusted, static mix levels maintained the dynamics intended by the performers.

For each STAAG configuration, the audio sample was played via the speakers, followed by a live male voice and clicks. For the voice and clicks, a male voice introduced each location for 30° increments from -90° to +90° by slating "location take color," where the "color" was red, orange, yellow, green, blue, indigo, or violet for each increment respectively. The same individual performed the clicks with a Petsmart TopPawTM dog-training clicker. This augmented the male voice, offering a robust cue with a consistent timbre.

The sound source capture utilized 27 different microphone positions to allow for different combinations of front/rear angles and spacings. Chart 2.1 shows the individual STAAG configurations. Nine different angular combinations were used, and repeated with three pairs of rear-facing microphones. In order to include cardioid, hyper-cardioid, and shotgun mics, Schoeps CMC6-MK4, Sennheiser MKH 8050, and Schoeps CMIT 5U shotgun mics were used, respectively. All microphone pads and filters were switched off, maintaining the best signal to noise ratio and emphasizing as much low-frequency energy in the ambiance as possible. It has been shown that this low-frequency content can contribute to a greater sense of LEV.8 A Millennia Media HV3-8ch stepped-gain microphone preamplifier, feeding an RME Fireface audio interface was employed to record each microphone. A laptop computer running a Magix Sequoia digital audio workstation played back and recorded all tracks simultaneously. Mic gains were adjusted and matched for each mic pair, and set to peak at roughly -6dBFS.

### TABLE 2.1 - Shows all mic types and combined pairs used for capture including STAAG configurations 1-27 and additional techniques recorded for comparisons. Only take # 1, 3, 11, 15, 18, 21, and 27 were used in subjective listening evaluations.

<table>
<thead>
<tr>
<th>Take #</th>
<th>Title</th>
<th>Front Mic Type</th>
<th>Front Angle</th>
<th>Rear Mic Type</th>
<th>Rear Mic Angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20-40 5U</td>
<td>8040</td>
<td>10</td>
<td>CMIT 5U</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>20-70 5U</td>
<td>8040</td>
<td>10</td>
<td>CMIT 5U</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>20-100 5U</td>
<td>8040</td>
<td>10</td>
<td>CMIT 5U</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>40-40 5U</td>
<td>8040</td>
<td>20</td>
<td>CMIT 5U</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>40-70 5U</td>
<td>8040</td>
<td>20</td>
<td>CMIT 5U</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>40-100 5U</td>
<td>8040</td>
<td>20</td>
<td>CMIT 5U</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>60-40 5U</td>
<td>8040</td>
<td>30</td>
<td>CMIT 5U</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>60-70 5U</td>
<td>8040</td>
<td>30</td>
<td>CMIT 5U</td>
<td>35</td>
</tr>
<tr>
<td>9</td>
<td>60-100 5U</td>
<td>8040</td>
<td>30</td>
<td>CMIT 5U</td>
<td>50</td>
</tr>
<tr>
<td>10</td>
<td>20-40 MK4</td>
<td>8040</td>
<td>10</td>
<td>CMC-6 MK4</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>20-70 MK4</td>
<td>8040</td>
<td>10</td>
<td>CMC-6 MK4</td>
<td>35</td>
</tr>
<tr>
<td>12</td>
<td>20-100 MK4</td>
<td>8040</td>
<td>10</td>
<td>CMC-6 MK4</td>
<td>50</td>
</tr>
<tr>
<td>13</td>
<td>40-40 MK4</td>
<td>8040</td>
<td>20</td>
<td>CMC-6 MK4</td>
<td>20</td>
</tr>
<tr>
<td>14</td>
<td>40-70 MK4</td>
<td>8040</td>
<td>20</td>
<td>CMC-6 MK4</td>
<td>35</td>
</tr>
<tr>
<td>15</td>
<td>40-100 MK4</td>
<td>8040</td>
<td>20</td>
<td>CMC-6 MK4</td>
<td>50</td>
</tr>
<tr>
<td>16</td>
<td>60-40 MK4</td>
<td>8040</td>
<td>30</td>
<td>CMC-6 MK4</td>
<td>20</td>
</tr>
<tr>
<td>17</td>
<td>60-70 MK4</td>
<td>8040</td>
<td>30</td>
<td>CMC-6 MK4</td>
<td>35</td>
</tr>
<tr>
<td>18</td>
<td>60-100 MK4</td>
<td>8040</td>
<td>30</td>
<td>CMC-6 MK4</td>
<td>50</td>
</tr>
<tr>
<td>19</td>
<td>20-40 MK4</td>
<td>8040</td>
<td>10</td>
<td>8050</td>
<td>20</td>
</tr>
<tr>
<td>20</td>
<td>20-70 MK4</td>
<td>8040</td>
<td>10</td>
<td>8050</td>
<td>35</td>
</tr>
<tr>
<td>21</td>
<td>20-100 MK4</td>
<td>8040</td>
<td>10</td>
<td>8050</td>
<td>50</td>
</tr>
<tr>
<td>22</td>
<td>40-40 MK4</td>
<td>8040</td>
<td>20</td>
<td>8050</td>
<td>20</td>
</tr>
<tr>
<td>23</td>
<td>40-70 MK4</td>
<td>8040</td>
<td>20</td>
<td>8050</td>
<td>35</td>
</tr>
<tr>
<td>24</td>
<td>40-100 MK4</td>
<td>8040</td>
<td>20</td>
<td>8050</td>
<td>50</td>
</tr>
<tr>
<td>25</td>
<td>60-40 MK4</td>
<td>8040</td>
<td>30</td>
<td>8050</td>
<td>20</td>
</tr>
<tr>
<td>26</td>
<td>60-70 MK4</td>
<td>8040</td>
<td>30</td>
<td>8050</td>
<td>35</td>
</tr>
<tr>
<td>27</td>
<td>60-100 MK4</td>
<td>8040</td>
<td>30</td>
<td>8050</td>
<td>50</td>
</tr>
</tbody>
</table>

After capturing the samples, the authors evaluated the 27 recordings and selected a set of seven samples to be evaluated by listeners (1, 3, 11, 15, 18, 21, and 27 from Chart 2.1). To reduce the number of total trials that listeners needed to evaluate, and therefore create a test that was a reasonable duration, samples were selected based upon the...
following criteria: width of presentation of ensemble from narrow to wide, ILEV from low to high, sense of correlated center image in both direct sound and late reflections, and at least two samples from each mic combination group. It should be noted that sample 21 was exchanged with 26 for investigations of ILEV. The characteristics of sample 21 seemed nondescript, and to the authors ear's, 26 showed more promise of capturing cues similar to those found in the AB-Omni samples.

2.3 Subjective Listening Tests

Nine volunteers between the ages of 20 and 33 years evaluated the chosen recordings with Sennheiser HD600 headphones fed by an RME Babyface audio interface and a PC computer. Since no audiogram was available, listeners were evaluated for their ability to localize sounds properly on the left and right, and then were asked to evaluate a monaural sample. If listeners did not localize this monaural sound to be within +/- 15°, their results were not used. Six listeners successfully met the criteria.

The test required completions of three sections. The first and the last involved a localization task, and the middle, an evaluation of ILEV. For the training at the beginning of sections one and two, listeners were read a script defining the task, followed by an evaluation of two training samples, always presented in the same order. The localization tasks contained samples of music followed by a male voice/clicks that included the front and rear facing mics. Once training was complete, subjects heard each of the seven samples (1, 3, 11, 15, 18, 21, and 26 from Table 2.1) in random order, and asked to provide feedback regarding the location of each instrument and click. Listeners were informed that there are no right or wrong answers, that sound sources are not necessarily symmetrical, and that they might localize sources to be inside or outside of their heads. Samples could be repeated, and there was no time constraint. Subjects required about 80 minutes on average to complete the test. The subject drew a series of symbols or letters (one for each instrument or click) on a chart showing a top-down cross-sectional view of a human head as shown in the top section of Figure 2.1.

![FIGURE 2.1 - Listener response graphic where top section shows localization form and bottom shows listener envelopment (ILEV) evaluation.](image)

The middle section of the test required the subject to evaluate ILEV. This was conducted in accordance with recommendations found in the ITU-R BS.1284 standard for subjective assessment of sound quality. Training began with listeners being read the definition of ILEV defined in the Background (Section 1.1) above. Since not all listeners were musicians, they were told additionally to focus their attention on judging the width or sense of being surrounded by the reverb. The samples contained short excerpts of the same music that was presented in the first section of the test. Each sample was presented by two anchor-point examples with the same content. The first anchor-point was monaural, and the second was a stereo example of a similar mic position with AB-Omni technique. Listeners were told to respond via the graphic (shown in the lower section of Figure 2.1) by placing an "X" anywhere on the numbered line. The anchor-points were described with the following text: "the first example will be a zero on the line, and the second example will be a ten. Respond by indicating how the third sample compares with the first two."

The last section of the test involved the same task as the first, and thus no training was necessary. The only difference here involved the exclusive presentation of the forward-facing microphone signals. No additional reverb
was added. This was done to compare the localization responses with the first section, and thus the influence of the rear-facing mics on the stereo image of the direct sounds.

3. RESULTS

After data collection, subjects' response graphs were organized by sample and analyzed for angular response or ILEV coefficients. Plots of the data for each of the three sections can be seen in order in Figures 3.1-3.3. Figures 3.1a and b show the mean localization points for the four quartet instruments and the voice/clicks respectively. Instruments on the extreme left and right (vln 1 and vlc respectively) localize to a range from ±65° to 80°. While the virtual ensemble on the stage only occupied about ±30°, it is not uncommon to have an accentuated width when listening via headphones. Since the available stereo field for headphones occupies the entire 180° range, listeners and engineers generally accept this exaggeration of reality, and find it desirable. Similar characteristics can be seen in Figure 3.3a, showing consistency between dry and ambient samples. ILEV ratings for all samples are greater than 6.3 (for 60-100 8050, with relatively high standard deviation), and at best 9.2 (for 20-100 5U). It should be noted that 20-100 5U not only shows the highest ILEV, but also the lowest standard deviation at 1.93.

![Figure 3.1](image-url)

**FIGURE 3.1** - (a: top-left) Localization responses for string quartet recordings with front and rear-facing mics at unity gain (b: top-right) responses for voice/click recordings with front and rear-facing mics at unity gain (c: bottom-left) average range for all responses to quartet for a given trial with standard deviations (d: bottom-right) average range for all responses to voice/clicks for a given trial with standard deviations
FIGURE 3.2 - Listener envelopment (ILEV) evaluations of STAAG configurations

FIGURE 3.3 - (a: top-left) Localization responses for string quartet recordings with front mics only
(b: top-right) responses for voice/click recordings with front mics only
(c: bottom-left) average range for all responses to quartet for a given trial with standard deviation
(d: bottom-right) average range for all responses to voice/clicks for a given trial with standard deviations
One of the primary goals of this study was to confirm the consistency of angular position of sound sources in the stereo image when using relatively narrow angles for forward-facing mics in semi-coincident techniques. The previous research showed average angular variation in localization responses for the extreme left/right sources (vln 1 and vlc for music, "red" and "violet" for voice/clicks) ranging as large as 40° for music and 35° for clicks. With the refinement of these angles, we now have responses ranging 14° for music and 9° for clicks. This is a promising result. With regard to sources at narrower angles, there are comparable variations in the music samples, and only about 5° of reduction in the voice/click samples, and further work should be done to minimize this variability.

The second intention of this study was to discover a configuration that enhanced ILEV by making use of highly directional polar patterns in the rear-facing mics. Two configurations from the microphone group utilizing the Schoeps CMIT 5U show promise. Between these two mics there is a direct correlation between rear-facing mic angle and ILEV, as well as a decrease in standard deviation. The fact that STAAG 20-100 5U shows an ILEV rating approaching that of an AB-Omni technique is also encouraging.

With regard to the best localizing configuration, STAAG 20-100 5U had the lowest standard deviation in every group. Combining this with the fact that it had the highest rating and second best standard deviation for ILEV, this microphone combination and configuration is clearly the best performer.

4. SUMMARY AND CONCLUSIONS

After referencing the collected data, it is apparent that there can be a wide range of results in the position of localization for the middle instruments and voice/clicks. The STAAG 20-100 5U configuration gave the most consistent localization points and convincing sense of ILEV. In all configurations the localization of sources was improved over the prior study by using front angles ranging 20° - 60°. This research confirms that angles much narrower than traditional stereo techniques such as ORTF and NOS can be used successfully and give reliable localization cues.

While it was expected that hyper-cardioid patterns for the rear facing mics would improve ILEV over cardioid, this proved not to be the case. These configurations showed lower envelopment than cardioid, and had no correlation between ILEV and mic angle. The high directivity provided by the implementation of shotgun mics is a clear advantage.

Several improvements can be made in specific areas. First, when integrating shotgun mics with the STAAG technique, angles wider than 100° should be considered. With regard to variation in localization of sound sources less than 45°, use of more highly-directional microphones and the implementation of an acoustic baffle may prove advantageous. Finally, listeners gave a wide range of responses in terms of perceived distance of the sound source. Off-axis coloration in the high frequencies of directional microphones can have a significant effect on distance perception, thus leading to varying results if microphones of different make/model are used. One should also consider investigations on how this overlap in pattern at high frequency affects distance perception.

ACKNOWLEDGMENTS

The authors would like to thank CIRMMT (Centre for Interdisciplinary Research in Music Media and Technology) for providing much of the recording equipment used for this study, and McGill University for providing the use of Redpath Hall. We would also like to thank Stuart Bremner for his assistance, positive attitude, and tireless energy during the overnight recording session, James Perrella for being an essential last-minute roadie, Aristides Rivas (cello) for his musical performance, and Jessica "The Explicatrix" Davis (violins I and II, viola) for her musical performances, editing, and multifaceted support of this research.

REFERENCES