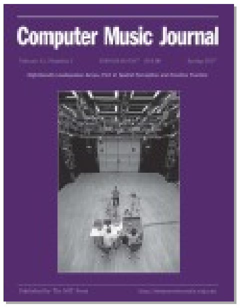
A Perceptual Investigation into Spatialization Techniques Used in Multichannel Electroacoustic Music for Envelopment and Engulfment



Hugh Lynch, Robert Sazdov

Computer Music Journal, Volume 41, Number 1, Spring 2017, pp. 13-33 (Article)

Published by The MIT Press

Hugh Lynch* and Robert Sazdov†
*Digital Media and Arts Research Centre
Department of Computer Science and
Information Systems
CSIS Building, University of Limerick
Limerick, Ireland
hughielynch@gmail.com
†School of Communication, Faculty of
Arts and Social Sciences
University of Technology Sydney
City Campus, 15 Broadway, Ultimo NSW
2007, Australia
robert.sazdov@uts.edu.au

A Perceptual Investigation into Spatialization Techniques Used in Multichannel Electroacoustic Music for Envelopment and Engulfment

Abstract: Composers of electroacoustic music have developed and creatively implemented various spatialization techniques for multichannel loudspeaker setups. What is not known is which of these spatialization techniques is most effective for exploiting the extended creative possibilities available in multidimensional sound. This article discusses an experiment investigating the perception of the spatial attributes of "envelopment" and "engulfment" within a high-density loudspeaker array. The spatialization techniques used in the experiment were timbre spatialization, spectral splitting, amplitude point-source panning, and dynamic spectral subband decorrelation. Three loudspeaker setups, or spatial dimensions, were investigated: horizontal-only; elevated-only; and three-dimensional, which consisted of both horizontal and elevated loudspeaker setups. Results suggest that dynamic spectral subband decorrelation was perceived as both the most enveloping and the most engulfing technique when compared to other techniques in these experimental loudspeaker configurations. We propose that the experimental results can be successfully implemented when composing electroacoustic music to exploit the creative possibilities in a high-density loudspeaker array or in other multichannel loudspeaker configurations.

Spatialization in Electroacoustic Music

Composers have engaged in multichannel loudspeaker diffusion since the earliest performances of electroacoustic music. In the early 1950s, Pierre Schaeffer and Pierre Henry used a four-channel configuration with an elevated loudspeaker included to spatialize their works (Harrison 1998). Composers such as John Cage, Karlheinz Stockhausen, Edgard Varèse, and Iannis Xenakis have spatialized works in various loudspeaker configurations, set up in concert halls or large rooms (Normandeau 2009). These configurations ranged from a four-channel system used by Schaeffer to a 55-loudspeaker system used by Stockhausen at the 1970 world's fair, Expo '70, in Osaka. Since the 1970s, standardized large-scale diffusion systems, such as the Acousmonium at the Maison de Radio in France (1974) and the Birmingham ElectroAcoustic Sound Theatre (1986) have been constructed. More recently, the Sonic Lab of the Sonic Arts Research Laboratory at Queen's University Belfast (2006), the Klangdom in the Kubus at the Center for Art and Media in Karlsruhe, Germany (2006), the György Ligeti Hall in the House for Music and Theater at the University of Music and Performing Arts Graz (2009), and the Spatial Music Workshop at Virginia Polytechnic Institute and State University (2014) have all set up high-density loudspeaker arrays that enable composers to work in three-dimensional space.

The preponderance of research in multichannel electroacoustic music has focused on theoretical approaches (e.g., Desantos, Roads, and Bayle 1997; Harrison 1998) or the practical implementation of diffusion or spatialization of sound (e.g., Chowning 1971). Composers have developed many processes, including granular, spectral, and panning techniques to diffuse and spatialize multichannel music (see Table 1).

It is evident that composers develop and use many techniques to spatialize multichannel sound, but it is unclear what perceptual effect or sonic

Computer Music Journal, 41:1, pp. 13–33, Spring 2017
doi:10.1162/COMJ_a_00401

Table 1. Multichannel Spatialization Techniques

Spatialization Technique	*Authors*
Spectral Approaches	
Frequency-domain processing	Torchia and Lippe (2004)
Spectral spatialization with boids	Kim-Boyle (2006)
Timbre spatialization	Normandeau (2009)
Spectral splitting	Wilson and Harrison (2010)
Equalator technique	Barreiro (2010)
Granulation Approaches	
Sound spatialization with particle systems	Kim-Boyle (2005)
Swarm lab	Davis and Rebelo (2005)
Spatial–swarm granulation	Wilson (2008)
Decorrelation and Panning Approaches	
Decorrelation	Kendall (1995)
Subband decorrelation	Potard and Burnett (2004)
Spectral-magnitude-based decorrelation	Wilson and Harrison (2010)
Pairwise and amplitude point-source panning	Chowning (1971)
Amplitude-dependent approach	Wilson and Harrison (2010)

Summary of techniques used by specific composers (for details cf. Lynch and Sazdov 2011a).

Table 2. Expressive Terms Used

Expressive Term	*Composers*
3-D musical space travel (Stockhausen 1971)	Karlheinz Stockhausen
Multiloudspeaker immersion (Emmerson 2007)	
Immersiveness (Normandeau 2009)	
Immersive reality (Rolfe 1999)	Chris Rolfe
Envelopment (Harrison 1999)	Jonty Harrison
Move within	Robert Normandeau
Immersed (Normandeau 2009)	
Envelope	Francis Dhomont
Soaked in sound (Basque and Watson 2004)	
Envelopes (Davis and Rebelo 2005)	Tom Davis
Localized effect	Scott Wilson and Jonty Harrison
Diffuseness (Wilson and Harrison 2010)	
Immersiveness (Peters, Marentakis, and McAdams 2011)	Various electroacoustic composers

Expressive terms used by composers to describe multichannel electroacoustic music (taken from Lynch and Sazdov 2011b).

experience, if any, is created by these approaches. Research by the authors (Lynch and Sazdov 2011b) revealed that composers have described some of their perceptual observations of multichannel music using expressive terms such as “enveloping” (Harrison 1999) and “immersive reality” (Rolfe 1999). Table 2 lists further examples.

The expressive terms commonly used by composers suggest that a perceptual sense of being enveloped or “surrounded” by sound is desirable when composing or listening to multichannel electroacoustic music. This would indicate that spatialization techniques for composing enveloping multichannel music are also desirable. Hence, the aim of this research is to formulate spatialization techniques that can be used to compose multichannel electroacoustic music that will be perceived by listeners as enveloping.

Kendall and Cabrera (2011) have stated that to make further advances in the practice of sound spatialization, a greater understanding of how listeners perceive multichannel sound is required. Robert Normandeau (2009) states that much research has been undertaken to explore the perception of distance and localization of noncomplex sound sources, but little is known about the perception of sound in space in a musical context. Normandeau goes on to argue that perceptual research experiments (e.g., Sazdov, Paine, and Stevens 2007) should be undertaken in parallel with music composition, to explore how space is perceived and what attributes of space are relevant in multichannel immersive environments. Can perceptual research in electroacoustic music be used to formulate enveloping spatialization techniques?

A significant portion of perceptual research in electroacoustic music consists of listener-response studies (e.g., Bridger 1989; Delalande 1989; Deliège 1989; Landy 1994; Smalley 1996). These studies are concerned with identifying pertinent sounding characteristics in works or defining types of listening behaviors that are used as listening strategies for the analyses of works. Other perceptual studies use approaches based on intention/reception studies, which are concerned with investigating the relationship between composers' compositional intentions and listeners' responses to presented works. The aim of these studies is to obtain both a wider and a deeper understanding or "meaning" of a work, and to identify whether there is a correlation between a composer's creative intentions and listeners' interpretations of a work (McCartney 2000; Weale 2005).

A number of perceptual evaluation experiments have investigated how electroacoustic music is perceived in a multichannel environment (Sazdov, Paine, and Stevens 2007; Adair, Alcorn, and Corrigan 2008; Sazdov 2011a, b). The studies incorporated research methodologies from the related disciplines of concert-hall acoustics and reproduced audio to investigate the effects of electroacoustic music on the perception of spatial attributes. Specifically, the studies investigate the effects of different frequency ranges, sonic complexities, and loudspeaker locations on the perception of the spatial attributes, spatial clarity, envelopment, and engulfment (the latter two terms are explained later in this article). Sazdov, Paine, and Stevens (2007) have suggested that novel compositional techniques can be formulated from the results of these perceptual studies.

Similar to the these earlier perceptual experiments by Sazdov and colleagues, the study presented here incorporates research methods from concert-hall acoustics and reproduced-audio research to investigate how listeners perceive spatial techniques in multichannel loudspeaker configurations. As mentioned earlier, research undertaken by the authors revealed that established composers describe their perceptual experiences of multichannel music using terms such as "envelopment." Hence, a number of the techniques listed in Table 1 were selected with the aim of determining which technique is perceived by listeners to be the most enveloping approach.

Perceptual Research in Related Disciplines

Spatial attributes are terms used in concert-hall and reproduced-audio research to describe spatial impressions or specific aspects of space (Nakayama et al. 1971; Barron and Marshall 1981). In concert-hall acoustics and reproduced-audio research, related physical measurements are used to quantify and predict spatial attributes rated by listeners (e.g., Bradley and Soulodre 1995; Berg and Rumsey 1999, 2001; Zacharov and Koivuniemi 2001; Soulodre, Lavoie, and Norcross 2002; Guastavino and Katz 2004). Research within these disciplines involves the rating of spatial attributes by participants presented with spatial audio scenes.

In the field of acoustics, many studies involve the perceptual evaluation of music in concert hall environments using spatial attributes (e.g., Barron 1971; Schroeder, Gottlob, and Siebrasse 1974; Barron and Marshall 1981; Beranek 1992; Bradley and Soulodre 1995; Kahle 1995; Morimoto, Iida, and Sakagami 2001). Concert-hall acoustics research is a discipline related to electroacoustic music, since multichannel performances commonly take place in large rooms or concert halls. In reproduced-audio research, studies have focused on the perceptual evaluation of reproduced audio, with findings concluding that specific attributes are important to the spatial sound quality in multichannel reproduction systems (e.g., Berg and Rumsey 1999, 2001, 2002; Zacharov and Koivuniemi 2001; Soulodre, Lavoie, and Norcross 2002). The current research adopts methodologies used in these related disciplines—namely, the rating of perceptual attributes to evaluate spatial techniques.

Spatial Attributes

As mentioned earlier, research has revealed that established composers have used expressive terms such as envelopment, immersiveness, or "surroundness" to describe their perceptual experience of multichannel electroacoustic music. This suggests that composers find these perceptual experiences desirable when composing or listening to multichannel sound. The spatial attributes chosen for the experiment were "ensemble envelopment" (Rumsey 1998) and "engulfment" (Sazdov, Paine, and Stevens 2007). These specific attributes were selected because they were deemed to describe the multidimensional nature of 3-D space, and because they were the expressive terms used by composers (as in Table 2), as well as having been previously used in perceptual experiments.

In concert-hall acoustics research, listener envelopment (LEV, see Barron and Marshall 1981) refers to late-lateral-arriving sound energy that creates an enveloping environment. Findings by Barron and Marshall (1981) and by Bradley and Soulodre (1995) reveal that lateral reflections are important in creating a perception of envelopment. A study by Morimoto, Iida, and Sakagami (2001) suggests that late arriving sound from the rear of the listening position increases the perception of envelopment. According to Berg and Rumsey (2001), envelopment created by sound arriving from all around the listener creates the perception of being immersed in a reverberant environment. In these studies, a sense of envelopment is perceived by late-lateral-arriving or rear-arriving reflections. This indicates that LEV is primarily perceived as a two-dimensional spatial attribute.

Ensemble envelopment, as defined by Rumsey (1998), is different from LEV. As discussed in concert-hall acoustics (cf. Barron and Marshall 1981), LEV is caused by late reverberant sound from the sides resulting in the feeling of envelopment. Rumsey states, however, that a sense of envelopment can be perceived by listening to a number of dry sources in a multichannel reproduction system where no late-arriving reflections are present. In later work, Rumsey states that this form of envelopment cannot be considered LEV in the traditional sense (Rumsey 2002). This specific definition is suitable for describing the perception of envelopment in multichannel arrays used by electroacoustic composers, because sound sources perceived within these arrays might not contain late reverberant energy, but instead might contain dry, direct sound sources. The studies discussed suggest that envelopment is perceived because of lateral or rear sounds, or it may be caused by sounds arriving from all directions around the listening position. Hence, we propose that the spatial attribute deemed most similar in definition to the terms expressed by composers is ensemble envelopment. From this point forward ensemble envelopment will simply be referred to as *envelopment*.

It is evident, from the first performance of electroacoustic music by Pierre Schaeffer and Pierre Henry in 1951, that those composers utilized multichannel diffusion systems with elevated loudspeakers included. Since then, many composers have formulated spatialization techniques and composed works for 3-D multichannel diffusion systems. This suggests that composers regard elevated sound as an important creative parameter when composing works. Sazdov, Paine, and Stevens (2007) have argued, however, that music composition has not adequately exploited the creative possibilities within 3-D multichannel loudspeaker configurations. Their investigation involved the perceptual evaluation of 3-D multichannel electroacoustic music. Findings from the experiment identified a unique elevated spatial attribute, engulfment. Whereas envelopment is defined as a sense of being surrounded by sound, engulfment is defined as a sense of being "covered" by sound (Sazdov, Paine, and Stevens 2007; Sazdov 2011a, b). An example of a real-life, perceptual sense of engulfment would be listening to thunder or fireworks exploding overhead from a listening position. Sazdov, Paine, and Stevens argue that these experiences do not envelop listeners, but "cover" them with sound. Sazdov (2015) argues, further, that elevated sound does not contribute to the perception of envelopment and is perceptually different from being surrounded by sound. In addition, the inclusion of this attribute can be used to further explore the creative possibilities of composing for configurations that include elevated loudspeakers.

It should be stated that other spatial attributes used in concert-hall acoustics and reproduced-audio research were considered for inclusion in our experiment. For example, apparent source width (Morimoto 1997), an attribute utilized in perceptual experiments (e.g., Potard and Burnett 2004), refers to the size or width that a source image inhabits and is most often associated with the front-arriving sound in a concert hall. Spatial impression was considered, as it is often used in perceptual experiments (e.g., Barron and Marshall 1981; Bradley 1994; Morimoto, Iida, and Sakagami 2001). We found that spatial impression was most often used to describe the perception of the entire spatial scene (a term defined more clearly in the following section). Spatial impressions were used in experiments to define the overall global spatial precepts in 2-D and, more recently, in 3-D space (Lee and Gribben 2014; Sazdov 2015). After further consideration, we concluded, however, that envelopment and engulfment were most relevant for inclusion in this study, because they are more descriptive in nature.

Perceptual Experiment

This section details the design of a perceptual experiment that aims to examine the perception of two spatial attributes in four spatial techniques presented in three multichannel loudspeaker configurations. This includes an outline of the experiment's aims, specifications of the loudspeaker configurations used, details of listeners who participated in the experiment, introduction and definition of the term "spatial scenes," the spatial diffusion used in the experiment, the equipment used and how the equipment was calibrated, the four spatial techniques under examination, and the procedure of the listening test.

Experimental Aims

The experiment was designed to examine the perception of the spatial attributes envelopment and engulfment in certain spatial techniques used in multichannel electroacoustic music. A listening experiment was run in the Spatialization and Auditory Display Environment (SpADE) at the Digital Media and Arts Centre (DMARC), University of Limerick. The aims of the experiment were to determine whether participants reported a perceived significant difference (1) between spatial techniques for envelopment and engulfment, and (2) between horizontal, elevated, and 3-D spatial scenes for envelopment and engulfment.

Experimental Loudspeaker Configuration

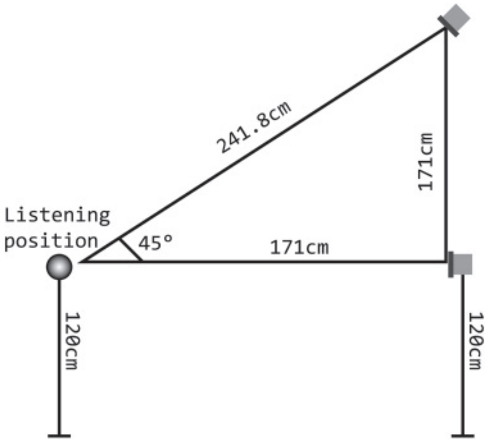
The experiment was run in SpADE, which consists of 32 Genelec 8030a loudspeakers with 14 loudspeaker configurations incorporated into the overall configuration, of which seven were horizontal-only and seven were 3-D configurations. SpADE provides the horizontal-only configurations monophonic, stereo, quadrophonic, six- and eight-channel, and the 5.1 and 7.1 configurations defined by the International Telecommunication Union. Silzle, Sunish, and Bachmann (2011) define a 3-D audio system as one that renders sound using both horizontal and elevated loudspeakers. Hence, SpADE also incorporates 3-D configurations proposed for film, such as Auro-3-D 9.1 and 10.2 (Holman 2008), Samsung 10.2 and 11.2 (Kim, Lee, and Pulkki 2010), and NHK 22.2 (Hamasaki, Hiyama, and Okumura 2005). In addition, SpADE's configuration includes an 8+8 3-D configuration: eight loudspeakers in the horizontal plane and eight elevated speakers.

Scott Wyatt (1999) states that electroacoustic composers commonly use eight-channel loudspeaker configurations. The eight-channel configuration can be configured as either a circular pattern or a pairs pattern. The circular pattern is effective if a composer is primarily working with monophonic source material. In the context of an amplitude pairwise-panning spatial technique, such as amplitude point-source panning (Chowning 1971), a monophonic source can be perceived as moving between two loudspeakers in the array. Because monophonic sources are primarily used in this instance, the presentation of stereo images is not a concern for this study. The pairs pattern is designed to be more effective for the presentation of stereo sound sources, with a primary concern being the maintenance of a stereo image within a listening space. The spatial techniques evaluated in the current experiment use monophonic sound sources within their spatial processes. For this reason, the techniques were evaluated in loudspeaker arrays having eight-channel configurations with circular patterns.

Three loudspeaker configurations within SpADE's configuration were used in this experiment: an 8-channel horizontal-only, an 8-channel elevated-only, and a 16-channel 3-D configuration. The horizontal-only configuration consisted of eight loudspeakers positioned equidistantly in a circular pattern with a radius of 1.71 m around a "sweet spot." The loudspeakers were positioned at 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315° in relation to the center sweet-spot position, at a height of 120 cm from the floor. The elevated-only configuration consisted of eight loudspeakers positioned equidistantly in a circular manner around the sweet spot, identical to the horizontal-only configuration but with a distance of 241.8 cm, because the loudspeakers were located at a height of 291 cm and with a 45°angle of elevation (see Figure 1). The elevated loudspeakers are positioned facing the listening position at a 45° angle. The 3-D configuration consisted of 16 loudspeakers, a combination of the horizontal- and elevated-only configurations as outlined here.

When spatial scenes were presented in the 3-D loudspeaker configuration, a delay was applied to all horizontal loudspeakers to compensate for the difference in distance from the sweet spot of the horizontal and the elevated loudspeakers. A delay of 2.08 msec (92 samples at 44.1 kHz) was applied equally to all horizontal loudspeakers within this configuration. This was achieved using a delay plug-in in Logic Pro 9.

Figure 1. Angular and distance relationship between elevated and horizontal loudspeakers (from Ronan, Piggott, and Sazdov 2012).



Participants

Twenty-four participants took part in the listening experiment. The participants included twelve undergraduate and twelve postgraduate students in the Bachelor of Science in Music, Media, and Performance Technology, and Master of Science in Music Technology programs at DMARC. Sixteen were men and eight were women, with a mean age of 26.3 years (ranging from 19 to 43 years of age). All participants reported that they had normal hearing ability, with five reporting to have undertaken a listening test in the past. The participants had been engaged in music or sound studies for between 1 and 20 years. All participants were deemed either experienced or expert listeners.

Seventeen participants reported that they were composers of music. Of these composers, nine reported that they had engaged in multichannel composition, but no participants reported that they had composed music for 3-D loudspeaker configurations. Furthermore, 13 participants stated that they listened to electroacoustic music, and 7 reported that they had attended a live concert of electroacoustic music.

All participants received course credit in exchange for taking part in the listening experiment.

Spatial Scenes

In this study, the term *spatial scene* refers to a composed sound source, or stimulus, that was spatialized using one of four techniques: timbre spatialization (Normandeau 2009), spectral splitting (Wilson and Harrison 2010), amplitude point-source panning (Chowning 1971), and dynamic spectral subband decorrelation (which will be described in more detail presently). These techniques were presented in three different multichannel loudspeaker configurations for perceptual evaluation.

The sound material chosen to create the stimulus was an excerpt from Claude Debussy's *Prélude à l'après-midi d'un faune* (1894). This excerpt was chosen because of its rich harmonic content. We applied digital signal processing (DSP) techniques to the excerpt in order to calibrate the stimuli and simulate sonic characteristics commonly heard in electroacoustic music. The DSP processes applied were time stretching and pitch transposition. The excerpt was first time-stretched by a factor of six. The stretched excerpt was transposed upwards or downwards by one or more octaves, so that most of the energy would lie in one of five frequency bands, creating five stimuli. The ranges of the frequency bands were 20–250 Hz, 250–2,000 Hz, 2–4 kHz, 4–8 kHz, and 8–20 kHz. Each stimulus was 15 seconds long, had 100-msec onsets and offsets applied, and had the same amplitude envelope. To remove any perceived loudness differences between them, each stimulus was presented from a single loudspeaker standing at a height of 120 cm, and the amplitude level for each excerpt was calibrated with a Acoustilyzer AL1 SPL meter to read 68 dB SPL (±0.1 dB), using a C-weighting filter and with time response set to slow. The SPL meter was located directly in front of the speaker at a distance of 171 cm and a height of 120 cm. This standardization eliminated participants' possible bias towards one excerpt over another based on perceived loudness. The levels were adjusted via the output bus in Logic Pro 9, resulting in each excerpt having the same loudness. The five manipulated excerpts were then combined to create one sonic layer, which was used as the stimulus for the study. All manipulations of the excerpt, and the final sonic layer, were exported as 24-bit, 44.1-kHz AIFF files.

We also undertook further calibration of the amplitude level of the spatial scenes when spatialized, using each of the four spatialization techniques in each loudspeaker configuration. For each technique, the overall amplitude level of the spatial scenes was calibrated within the horizontal-only, elevated-only, and 3-D loudspeaker configurations with the same SPL meter settings as before (68 ± 0.1 dB SPL, C weighting, and slow response). For each technique, the stimulus was routed to all loudspeakers in the configuration, and the calibration was undertaken using the AL1 SPL meter located at the center listening position in each loudspeaker configuration (cf. Figure 1).

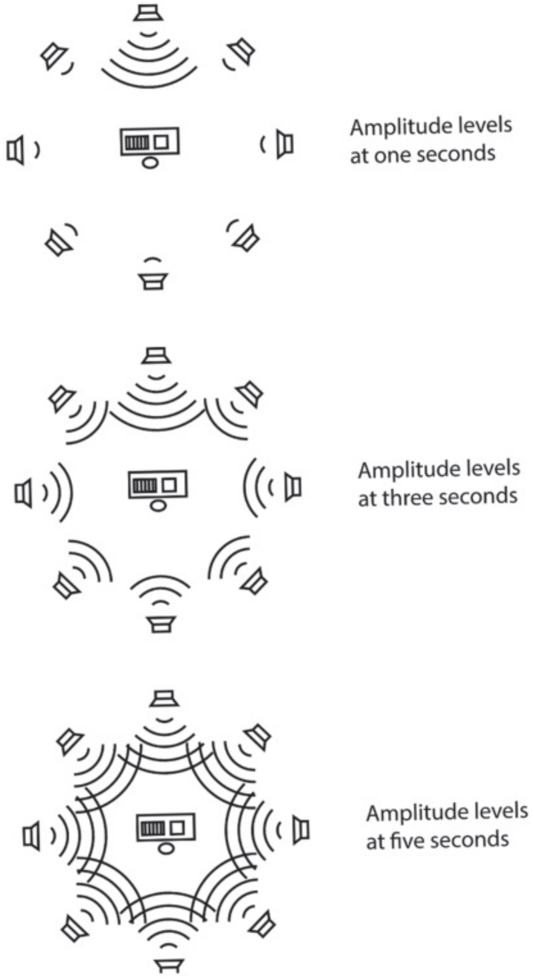
Spatial Diffusion

As noted by Cheng and Wakefield (2001), composers commonly use sound diffusion in multichannel electroacoustic music. Sound diffusion refers to the practice of localizing and moving sound throughout a space using multiple loudspeakers. This can be achieved by controlling the amplitude levels, equalization, and placement of sound sources. A sound diffusion process was applied to each spatialization technique presented in the horizontal-only, elevated-only, and 3-D configurations. For each spatial technique, diffusion was integrated at the start of a presented spatial scene by adjusting the amplitude levels of individual loudspeakers. This was done so that the sound source was initially perceived as originating from a frontal location only and then, over the course of five seconds, gradually became audible in each speaker within the configuration (see Figure 2).

Equipment and Calibration

Three different loudspeaker configurations were constructed within one setup. These consisted of an

Figure 2. The amplitude levels of each loudspeaker was adjusted to create a sound diffusion for each spatial scene.



8-channel horizontal-only, an 8-channel elevated-only, and a 16-channel 3-D configuration. The specifications of each configuration were outlined in the previous section. All 16 loudspeakers were identical Genelec 8030a self-powered speakers. An Apple Mac Pro running Logic Pro 9 was used to present the experiment. The Mac Pro was connected via Firewire to an M-Audio Audiophile 192 PCI interface. The 16 analog outputs of the interface were directed to the 16 loudspeakers used. These outputs were directly connected to the 8030a loudspeaker inputs using ¼-in XLR connectors. All loudspeakers in SpADE were calibrated with pink noise, using the AL1 SPL meter with the same settings previously specified (C-weighting and slow response) at a distance of 1 m. The levels were adjusted via the output buses in Logic Pro so that each loudspeaker read 70 dB SPL (±0.1 dB).

Spatial Techniques

We reviewed spatial techniques used by composers and identified, through practical implementation and perceptual observations, that a number of techniques were perceived as more enveloping or engulfing than others (see Lynch and Sazdov 2011b). The first three spatialization techniques selected for the experiment were timbre spatialization (TS), spectral splitting (SS), and amplitude point-source panning (APSP), which we now describe briefly. A fourth technique used in the experiment, dynamic spectral subband decorrelation (DSSD), is described in greater detail later in this article.

Timbre Spatialization

Timbre spatialization was implemented by assigning a dynamic band-pass filter to each loudspeaker within a configuration. Four different band-pass filters were used. Each band-pass filter was programmed so that all frequencies of the input sound source were filtered out, except for the specified frequency range. These ranges were labeled low (20–250 Hz), low midrange (250–2,000 Hz), high midrange (2–4 kHz), and high (4 kHz and higher). These ranges were chosen so that all frequencies within the audible frequency range, although fragmented, were presented for each spatial scene. The dynamic element of the technique consisted of changing the bandwidth of each band-pass filter over the duration of a scene. The opening (or widening) of each filter's bandwidth started after 5 seconds and gradually increased until the end of each spatial scene. At that point each filter was completely open, with the full spectrum of each sound source left unfiltered.

For each configuration, the four band-pass filters were assigned to the loudspeakers in consecutive order from low to high. For example, in the horizontal configuration the low filter was assigned to the loudspeaker positioned at 0°, the low midrange filter was assigned to the loudspeaker positioned at 45°, the high midrange filter was assigned to the loudspeaker positioned at 90°, and the high filter was assigned to the loudspeaker positioned at 135°, with this filter assignment sequence repeating for the remaining loudspeakers positioned at 180°, 225°, 270°, and 315°. The same mapping of band-pass filters to loudspeaker positions was used both in the elevated-only and in the 3-D configurations.

Timbre spatialization was perceived as both more enveloping and more engulfing then other techniques reviewed. One possible reason for this could be that the technique includes lateral and rear sound, which, as noted previously, contributes to a perception of envelopment. Additionally, TS contains dynamic filtering spatial processes that result in the perception of sound moving through and around the listening space. It is suggested that this perceived movement of sound contributes to an increased perception of envelopment and engulfment in comparison to spatial techniques where sound is not perceived as moving.

Spectral Splitting

Spectral splitting is based on varying frequency responses, proximity, and orientation of elevated loudspeakers in a heterogenous loudspeaker system (Wilson and Harrison 2010). The technique was applied to the 3-D configuration by routing the sound source to each loudspeaker. Each elevated loudspeaker in the configuration was assigned a high-pass filter programmed to cut off all frequencies below 1,000 Hz. No high-pass filters were assigned to the horizontal loudspeakers in this configuration.

The application of the technique in the elevated-only configuration is the same application as used in the elevated-only portion of the 3-D configuration described earlier. Conversely, the application of the technique in the horizontal-only configuration is the same application as is used in the horizontal portion of the 3-D configuration described earlier. Because the horizontal-only configuration does not contain any elevated loudspeakers, no high-pass filters were assigned to loudspeakers in this configuration. Spectral splitting was developed specifically for multichannel configurations with elevated loudspeakers included.

Amplitude Point-Source Panning

Amplitude point-source panning consists of placing monophonic sound signals on individual loudspeakers within an arbitrary loudspeaker configuration. Perceived movement of sounds between loudspeakers is achieved by changing the amplitude levels of individual loudspeakers within a configuration.

In this experiment, this technique was applied by directly assigning monophonic sound sources to each individual loudspeaker within a configuration. The amplitude levels of all loudspeakers were globally controlled over the spatial scene's 15-second duration. In the first 5 seconds of each spatial scene, the amplitude levels of each loudspeaker were automated to gradually increase from zero (i.e., no sound) to maximum amplitude, where they remained static.

When evaluated, the technique was perceived as more enveloping, because when sound is point sourced to each loudspeaker in an eight-channel loudspeaker configuration, sound is perceived as emanating from all directions within the listening space. This perceptual observation is consistent with findings from research in reproduced audio (Griesinger 1999; Morimoto, Iida, and Sakagami 2001; Bradley and Soulodre 1995) suggesting that lateral and rear sound contributes to the perception of listener envelopment.

Dynamic Spectral Subband Decorrelation

This section is a description of a novel spatialization method used in the perceptual experiment, called dynamic spectral subband decorrelation (DSSD). In this method, the sound source is first split into different frequency bands using band-pass filters. Each filtered source is routed to a speaker. Dynamic spectral and decorrelation processes are applied to each subband signal. The approach is implemented in Logic Studio by creating 16 audio channels within the software's Arrange window, routing a monophonic sound source to the input of each channel, and sending each signal output to a

Table 3. Settings for Dynamic Spectral Subband Decorrelation

Start Frequency Range	*End Frequency Range*	*Loudspeaker Location*
Low: 20–250 Hz	High: 4–20 kHz	0°
Low-Mid: 250–2,000 Hz	High-Mid: 2–4 kHz	45°
High-Mid: 2–4 kHz	Low-Mid: 250–2,000 Hz	90°
High: 4–20 kHz	Low: 20–250 Hz	135°
Low: 20–250 Hz	High: 4–20 kHz	180°
Low-Mid: 250–2,000 Hz	High-Mid: 2–4 kHz	225°
High-Mid: 2–4 kHz	Low-Mid: 250– 2,000 Hz	270°
High: 4–20 kHz	Low: 20–250 Hz	315°

The table shows the subband signals' band-pass filter start and end frequency ranges, and the loudspeaker locations. For each subband signal, a decorrelation delay of 10 msec was applied after 5 seconds and increased to 20 msec after 10 seconds. The filter sweep for all frequency ranges starts at 5 seconds and finishes at 15 seconds.

loudspeaker signal input in the array. Each audio channel is assigned a band-pass filter that is used to filter the channel output signals to a specific frequency range. Logic Studio's Channel EQ plug-in was used as the band-pass filter.

The frequency range of each filter consists of one of the same four ranges used in TS: low (20–250 Hz), low midrange (250–2,000 Hz), high midrange (2–4 kHz), and high (4–20 kHz). All of the band-pass filters have a roll-off of 48 dB per octave, resulting in a sharp cut in frequency content outside each of the filter's bandwidth ranges. The spatial scene has a duration of 15 seconds. After 5 seconds, each subband signal's filter bandwidth begins to move to a different frequency range. Specifically, all low filter frequency bands move to high, all high to low, all low midrange to high midrange, and all high midrange to low midrange. Table 3 summarizes each of the subband signal's filter starting and finishing ranges, and the loudspeaker location.

The movement of each subband signal's filter frequency band is implemented by changing the lower and upper cut-off frequency values in each filter. For example, the filter located at 0°, with an initial low frequency range of 20–250 Hz, moves to the high frequency range of 4–20 kHz. Its lower cut-off frequency of 20 Hz changes to 4 kHz and its upper cutoff frequency changes from 250 Hz to 20 kHz. The upper and lower cutoff values of the filters are programmed to change using Logic's Channel EQ plug-in. The changes in values are recorded using the automation function in Logic Studio. During spatialization, the lower and upper cutoff frequency values of the filters' frequency bands change at the same speed. The filters' Q and roll-off values do not change. This dynamic filtering process is similar to what is achieved when implementing a band-pass filter sweep. As each band-pass filter sweeps to a different frequency-band register, the spectral content of each subband signal changes.

The effect of this process is a perception that sound sources gradually move around the listening space. For example, the filtered source routed to the loudspeaker located at 0° changes from a low to a high frequency register, and the source routed to the speaker located at 135° changes from a high to a low frequency register. This leads to the perception that the low-filtered source moves from the front center to rear right and the high source moves from rear right to the front center of the listening space. The speed at which the filter's upper and lower cutoff frequencies move is dependent on the duration of the spatialization. The duration of the spatial scene is 15 seconds and the movement of the filter takes place after 5 seconds, meaning that it takes 10 seconds for each filter frequency range to move or sweep from its initial starting range to its new frequency range.

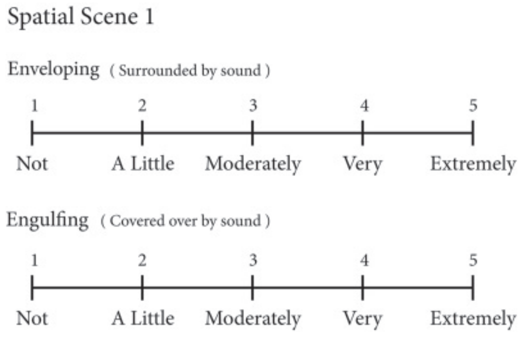


Figure 3. Participants rated each spatial scene using a 5-point Likert scale.

A decorrelated technique (Kendall 1995) is implemented by applying a temporal offset to each subband signal within the configuration. For each subband signal described above, an additional duplicated, decorrelated subband signal is sent to the same loudspeaker output. Decorrelation is implemented by applying a time delay to the duplicated signal, using Logic Studio's Sample Delay plug-in. The duration of the spatial scene is 15 seconds. For each decorrelated subband signal, the following time delay is applied for each spatial scene: at zero seconds there is no time delay (i.e., a delay of 0 msec), at 5 seconds a time delay of 10 msec is applied, and at 10 seconds the time delay increases to 20 msec.

Listening Test Procedure

The listening experiment was run individually for each participant. The experiment consisted of a training phase and an experiment phase. The training phase consisted of explaining the experiment procedures and providing a definition of each of the spatial attributes to be evaluated. Participants were informed that a series of spatial scenes would be presented and that they were to rate each excerpt perceptually for levels of envelopment and engulfment. The following explanation of terms was included in the information sheet for the participants:

Envelopment—the sense of being surrounded by sound.
Engulfment—the sense of being covered by sound.

A short verbal example of where each attribute could be perceived in a natural environment was also provided. These were "a sense of envelopment could be perceived when surrounded by a crowd of people," and "a sense of engulfment could be perceived from lightning or fireworks."

Each participant undertook a short training session prior to the listening experiment. Each was seated in the designated sweet spot, located in the center of the loudspeaker configuration. Six spatial scenes using two additional spatial techniques not used in the experiment were presented in a short training session. These techniques were "static" timbre spatialization (Normandeau 2009) and full-band decorrelation (Kendall 1995). A five-level Likert scale was used to elicit participant responses (see Figure 3). The procedure of rating spatial scenes using the Likert scale was explained to participants. They were asked to mark an "X" at any point along the scales, not just on the indicated anchor points.

This training session was undertaken to familiarize the participants with the format of the listening experiment. Participants were advised of the possible advantages of head movement (left–right and up–down movement) in perceiving sound, and they were advised that they could move their heads if they wished to do so. Participants were asked to rate each scene after it had finished playing. No other information was provided to the participants regarding the aims of the experiment. The experimental phase consisted of each participant rating 60 randomly ordered spatial scenes for perceived levels of envelopment and engulfment. The four spatial techniques were presented on the horizontal, elevated, and 3-D loudspeaker configurations. The twelve test conditions were repeated five times each, and the experiment lasted approximately 23 min. A 10-second break was allocated between each scene for the first 24 scenes, a 9-second break between each scene for the next 12, an 8-second break for the next 12, and a 7-second break for the remaining 12 scenes. The gradual shortening of breaks between scenes was applied because we found in the pilot study that participants did not need as long a break to complete the rating of each scene as the experiment progressed. Table 4 shows a list of the

Table 4. Variables Controlled in the Experiment

Controlled Variables	
Amplitude envelope	All spatial scenes had amplitude envelopes.
Sound pressure level	All spatial scenes were calibrated from the center listening position to have an identical, 68-dB SPL using C-weighting with response time set to slow.
Frequency ranges	Four frequency ranges (low, low midrange, high midrange, and high) were used to construct the stimulus used for all spatial scenes in the experiment, incorporating all frequencies between 20 Hz and 20 kHz.
Onset and offset times	The onset and offset for each spatial scene were both 100 msec.
Sonic complexity	The stimulus was manipulated using the DSP techniques of time stretching and transposition.
Stimuli length	All spatial scenes were 15 seconds in length, including onset and offset times.
Calibration of loudspeaker configuration	All loudspeakers in the configuration were calibrated using pink noise to measure 70 dB SPL at 1 meter, with C-weighting.
Participant seating	Each participant was positioned in the center listening position.
Spatial techniques	Timbre spatialization, spectral splitting, amplitude point-source panning, and dynamic spectral subband decorrelation were used.
Motion trajectory	The same motion trajectory was used for all presented spatial scenes.

Table 5. Ecological Variables not under Experimental Control

Relevant Ecological Variables	
Listening Room Environment	Experiment was run in a room similar to where multichannel electroacoustic music performance would be held.
Room Reverberation Time	$T_m = 0.34$ sec (Ronan, Piggott, and Sazdov 2012)
Complex Stimuli	The spatial scenes were manipulations of orchestral music, *Prélude àl'après-midi d'un faune* (1894), by Claude Debussy.
Stimuli Manipulation	The manipulations of the excerpt were typical of electroacoustic music, i.e. time stretching and transposition.
Amplitude Envelope	The amplitude envelope was not constant or fixed; it varied throughout the duration of each spatial scene. The amplitude envelope was identical for all spatial scenes, however.

controlled variables within the experiment design. Lists of the variables within the experiment that are considered ecologically valid are shown in Table 5.

Experiment Results

The experiment is concerned with examining the effects of sound-spatialization techniques on participants' perceptual ratings of spatial attributes. The spatial techniques were presented in different loudspeaker configurations. Specifically, the aim of the experiment was to identify which technique was perceived by participants as the most enveloping and which the most engulfing. To review: The spatial techniques under evaluation were APSP, SS, TS, and DSSD, and the loudspeaker configurations were horizontal-only, elevated-only, and 3-D. Hence, the independent variables are spatial techniques and loudspeaker configurations, and the dependent variables are the two spatial attributes envelopment and engulfment.

A factorial analysis of variance (ANOVA) was used to examine the influence of the two independent variables on the dependent variables. The analysis was used to identify a main effect of contributions for each independent variable and whether there was a significant interaction effect between the independent variables. A main effect is seen when an independent variable affects or influences the rating value of a dependent variable. An interaction effect, on the other hand, is seen when an interaction between two independent variables influences the rating value of the dependent variable. In statistical analysis, the null hypothesis refers to a statement or default position that there is no relationship between two factors. In significance testing, a null hypothesis is rejected because of data that is found to be significant, but not accepted or proved. There are three null hypotheses in this two-way ANOVA:

1. Participants' spatial attribute perceptual ratings for loudspeaker setups are equal.
2. Participants' spatial attribute perceptual ratings for techniques are equal.
3. There is no interaction or relationship between the spatial techniques and loudspeaker configurations.

To test Null Hypotheses 1 and 2, we analyzed the main effect of each independent variable's spatial attribute ratings. To test Null Hypothesis 3, we tested the interaction effect between independent variables. The main effect involves an independent variable's effect on the dependent variable value, with any interaction between independent variables being ignored. For example, the effect of spatial techniques on envelopment ratings is presented with the influence of loudspeaker configuration being ignored. As mentioned, the interaction effect is the effect of one independent variable on another independent variable. For example, the effect of spatial techniques on listener envelopment ratings is dependent on loudspeaker configuration.

Further, a pairwise comparison analysis was undertaken in the form of a Bonferroni post hoc test for each independent variable. The post hoc test is used to determine which pairs of means in the spatial-technique ratings for envelopment and engulfment are significantly different. In addition, a separate post hoc test is used to determine which pairs of means within the loudspeaker configuration ratings for envelopment and engulfment are significantly different. The post hoc tests are used to identify which technique and which loudspeaker configuration are rated significantly higher for envelopment and engulfment. Using a pairwise comparison analysis, we present the mean values for the envelopment and engulfment ratings of the spatial techniques in relation to loudspeaker configurations. This provides an indication of each spatial technique's highest rating for envelopment and engulfment. It also indicates in which specific loudspeaker configuration this highest rating is perceived.

Table 6. Factorial ANOVA Results

Envelopment				
Effect	*dof*	*F*	*p*	η_p^2
Main Effect: Loudspeaker	2	2.498	0.032	0.024
Main Effect: Techniques	3	12.810	0.00	0.119
Interaction Effect: Techniques–Loudspeakers	6	0.415	0.869	0.009

ANOVA results for the perception of envelopment. For each effect, the degrees of freedom (dof) and values for the F*-test (*F*),* p*, and eta-squared (η_p^2) are shown.*

In statistical significance testing, the p-value is the probability of acquiring a test statistic result. A null hypothesis is rejected when the p-value is found to be less than a predetermined significant level. When the p-value is over this level, the hypothesis is not rejected. Within this testing, the predetermined significance level is 0.05, so data with $p < 0.05$ is reported as significant. The degrees of freedom are the number of values in a final calculation of a statistic that are free to vary. The F-test assesses the hypothesis that the means of a set of normally distributed populations are equal if they have the same standard deviations. Eta-squared (η_p^2) is the measure of effect size for use in an ANOVA.

Envelopment Results

A factorial ANOVA was undertaken to find the main effects for loudspeaker configurations and spatialization techniques, and the interaction effect between configurations and techniques, for the perception of envelopment (see Table 6).

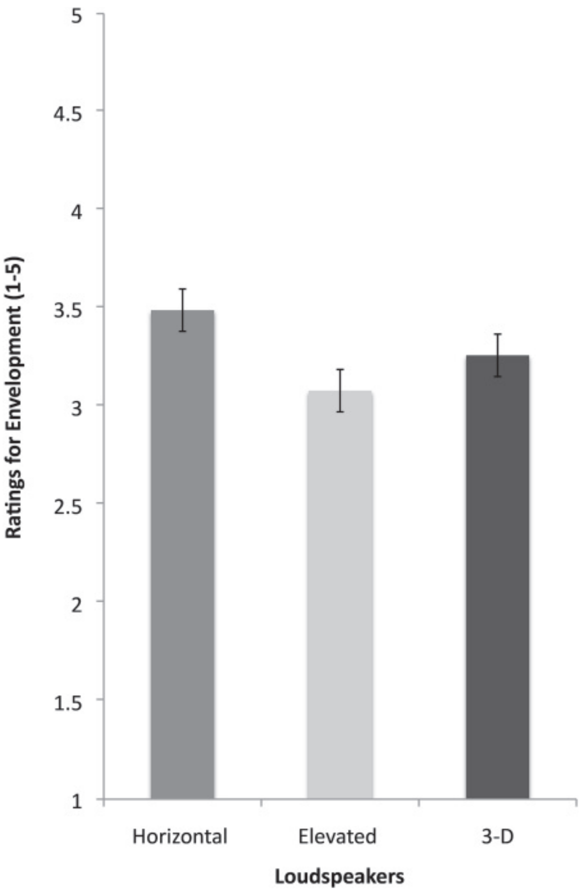


Figure 4. Post hoc comparison for the perception of envelopment for horizontal-only, elevated-only, and 3-D loudspeaker configurations. Error bars indicate standard error of means.

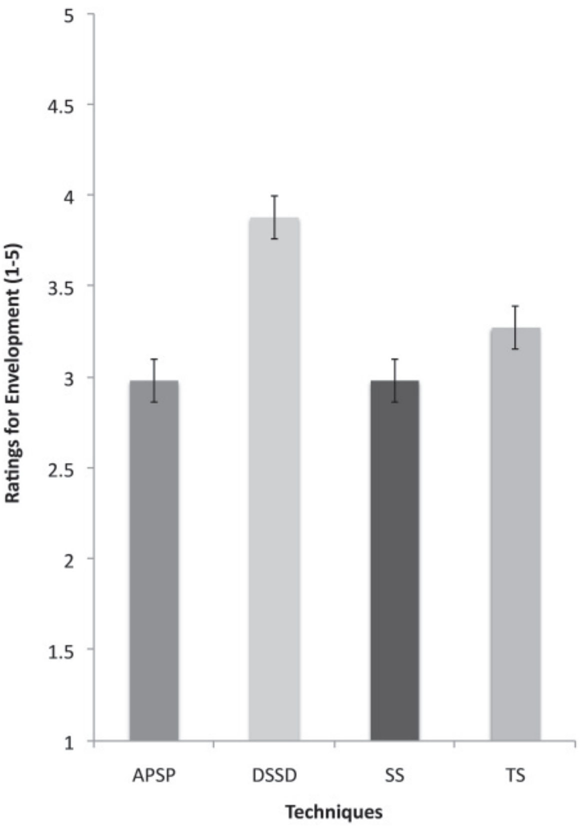


Figure 5. Post hoc comparison for the perception of envelopment for amplitude point-source panning (APSP), dynamic spectral subband decorrelation (DSSD), spectral splitting (SS), and timbre spatialization (TS).

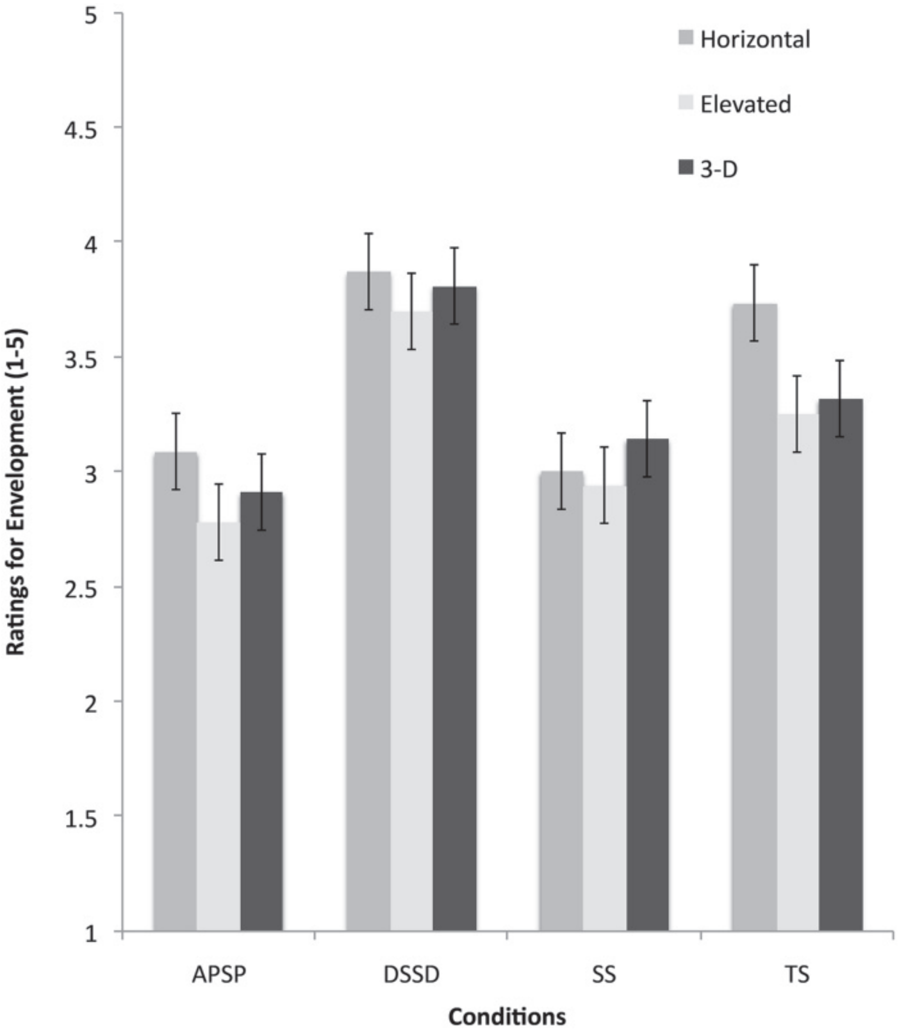
The main effect for loudspeakers, $F(2, 285) = 2.498$, $p < 0.05$ indicates a significant difference between horizontal-only (mean $\bar{x} = 3.49$, standard deviation $s = 0.95$), elevated-only ($\bar{x} = 3.08$, $s = 1.18$), and 3-D ($\bar{x} = 3.26$, $s = 1.00$) configurations (see Figure 4). As shown in Figure 5, the main effect for techniques $F(3, 284) = 12.810$, $p < 0.05$ suggests a significant difference between APSP ($\bar{x} = 2.98$, $s = 1.06$), DSSD ($\bar{x} = 3.87$, $s = 0.92$), SS ($\bar{x} = 2.98$, $s = 1.04$), and TS ($\bar{x} = 3.27$, $s = 0.97$). A Bonferroni post hoc comparison was undertaken to determine which pairs of means were significantly different for spatial techniques and loudspeakers.

The post hoc comparison results for loudspeaker configurations reveals a significant difference between horizontal-only ($\bar{x} = 3.48$, $s = 0.95$) and elevated-only ($\bar{x} = 3.08$, $s = 1.18$) configurations, $p < 0.05$. There was no significant difference between horizontal-only ($\bar{x} = 3.48$, $s = 0.95$), and 3-D ($\bar{x} = 3.26$, $s = 1.00$) configurations, $p = 0.135$. In addition, there was no significant difference between elevated-only ($\bar{x} = 3.08$, $s = 1.18$) and 3-D ($\bar{x} = 3.26$, $s = 1.00$) configurations, $p = 0.257$.

The post hoc comparisons of spatial techniques, shown in Figure 5, indicate that there are significant differences between DSSD ($\bar{x} = 3.87$, $s = 0.92$) and APSP ($\bar{x} = 2.98$, $s = 1.06$), between DSSD ($\bar{x} = 3.7$, $s = 0.62$) and SS $\bar{x} = 3.0$ ($s = 0.69$), and between DSSD ($\bar{x} = 3.87$, $s = 0.92$) and TS ($\bar{x} = 3.27$, $s = 0.97$), $p = < 0.05$.

The interaction effect between techniques and loudspeakers was not significant $F(6, 276) = 0.415$, $p = 0.869$ (see Figure 6). This result suggests that loudspeaker configurations used did not significantly affect envelopment ratings for techniques.

Figure 6. Participants' ratings for the perception of envelopment for all spatial techniques and loudspeaker configurations.



A post hoc comparison was run for multiple comparisons to discover which level of interaction between loudspeaker and spatialization techniques is significant. Amplitude point-source panning was not significantly different between horizontal-only ($\bar{x} = 3.2$, $s = 0.91$) and elevated-only ($\bar{x} = 2.8$, $s = 0.98$), p = 0.15, nor between horizontal-only ($\bar{x} = 3.2$, $s = 0.91$) and 3-D ($\bar{x} = 2.8$, $s = 1.25$), configurations, $p = 0.220$. Dynamic spectral subband decorrelation was not significantly different between horizontal-only ($\bar{x} = 4.0$, $s = 0.77$) and elevated-only ($\bar{x} = 3.7$, $s = 1.23$) configurations, $p = 0.249$, nor was there a significant difference between horizontal-only ($\bar{x} = 4.0$, $s = 0.77$) and 3-D ($\bar{x} = 3.8$, $s = 0.69$) configurations, $p = 0.525$. Spectral splitting was not significantly different between horizontal-only ($\bar{x} = 3.0$, $s = 1.04$) and elevated-only ($\bar{x} = 2.9$, $s = 1.07$) configurations, $p = 0.665$, nor between horizontal-only ($\bar{x} = 3.0$, $s = 1.04$) and 3-D ($\bar{x} = 2.9$, $s = 0.70$) configurations, p = 0.430. Timbre spatialization was significantly different between horizontal-only ($\bar{x} = 3.5$, $s = 0.80$) and elevated-only ($\bar{x} = 2.8$, $s = 1.25$) configurations, $p < 0.05$ and between elevated-only ($\bar{x} = 2.8$, $s = 1.25$) and 3-D ($\bar{x} = 3.3$, $s = 0.44$)

Table 7. Factorial ANOVA Results for Perception of Engulfment

Engulfment				
Effect	*df*	*F*	*p*	η_p^2
Main Effect: Loudspeaker	2	0.846	0.430	0.006
Main Effect: Techniques	3	5.180	0.002	0.052
Interaction Effect: Techniques × Loudspeakers	6	3.047	0.007	0.62

For each effect, the degrees of freedom (dof) and values for the F*-test (*F*),* p*, and eta-squared (η_p^2) are shown.*

configurations, $p < 0.05$. There was no significant difference between horizontal-only ($\bar{x} = 3.5$, $s = 0.80$) and 3-D ($\bar{x} = 3.3$, $s = 0.44$), $p = 0.387$ configurations.

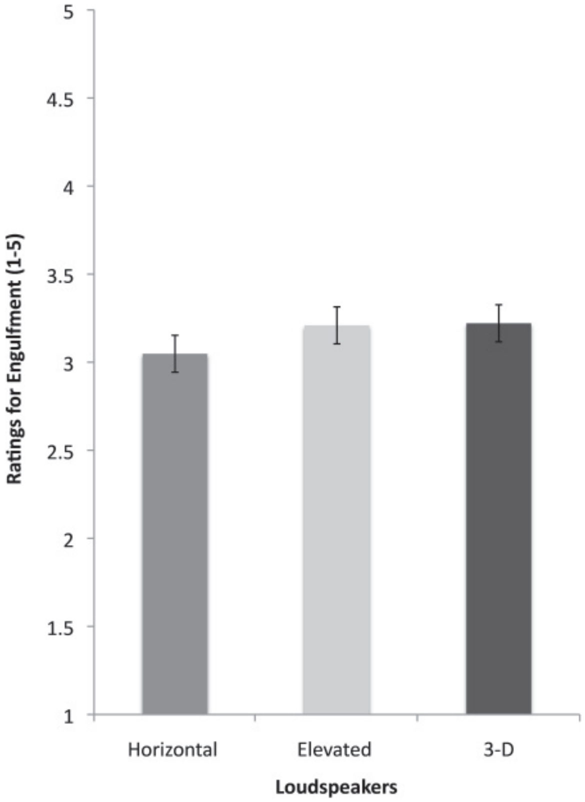
Engulfment Results

We undertook a factorial ANOVA to find the main effect for techniques, loudspeakers, and the interaction effect between the four spatialization techniques and the three loudspeaker configurations for the perception of engulfment (see Table 7).

The main effect for loudspeakers $F(2, 285) = 0.846$, $p = 0.430$, was not significantly different between horizontal-only, elevated-only, and 3-D configurations (see Figure 7). As shown in Figure 8, the main effect for techniques $F(3, 284) = 5.180$, $p < 0.05$, suggests a significant difference between APSP ($\bar{x} = 2.92$, $s = 0.88$), DSSD ($\bar{x} = 3.5$, $s = 1.04$), SS ($\bar{x} = 3.1$, $s = 1.04$) and TS ($\bar{x} = 3.0$, $s = 1.04$). The interaction effect between techniques and loudspeakers was significant, $F(6, 276) = 3.047$, $p < 0.05$ (Figure 9).

The post hoc comparison results for loudspeaker configurations reveals no significant difference between horizontal-only ($\bar{x} = 3.0$, $s = 0.97$) and elevated-only ($\bar{x} = 3.2$, $s = 1.03$) configurations, $p = 0.838$. There was no significant difference between horizontal-only ($\bar{x} = 3.0$, $s = 0.97$) and 3-D ($\bar{x} = 3.22$, $s = 1.07$) configurations, $p = 0.735$. In addition, there was no significant difference between elevated-only ($\bar{x} = 3.2$, $s = 1.03$) and 3-D ($\bar{x} = 3.22$, $s = 1.07$) configurations, $p = 1.000$. This suggests that Null Hypothesis 1 can not be rejected.

Figure 7. Post hoc comparison for the perception of engulfment for horizontal-only, elevated-only, and 3-D loudspeaker configurations.

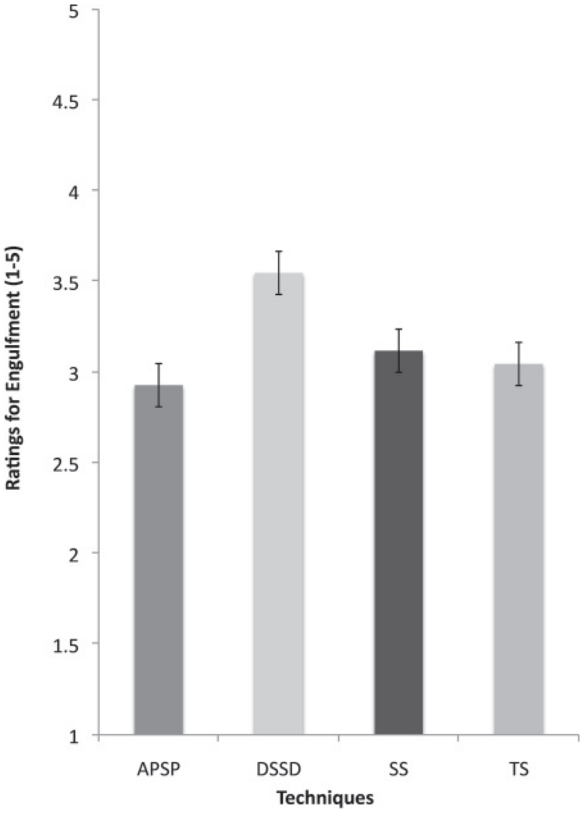


The post hoc comparison results for spatial techniques indicate that there was a significant difference between DSSD ($\bar{x} = 3.5$, $s = 1.04$) and APSP ($\bar{x} = 2.92$, $s = 0.88$), between DSSD ($\bar{x} = 3.5$, $s = 1.04$) and SS ($\bar{x} = 3.1$, $s = 1.04$) and between DSSD ($\bar{x} = 3.5$, $s = 1.04$) and TS ($\bar{x} = 3.0$, $s = 1.04$), with $p < 0.05$ in all cases.

The interaction effect between loudspeakers and techniques showed $F(6, 276) = 3.047$, $p < 0.05$. This suggests that the effect of a technique's engulfment rating is affected by a loudspeaker configuration.

A post hoc comparison was run for multiple comparisons to discover which levels of interaction between loudspeaker and techniques were significant. Amplitude point-source panning was not significantly different between horizontal-only ($\bar{x} = 2.8$, $s = 0.94$) and elevated-only ($\bar{x} = 3.0$, $s = 0.97$) configurations, $p = 0.465$, nor between horizontal-only ($\bar{x} = 2.8$, $s = 0.94$) and 3-D ($\bar{x} = 2.8$, $s = 0.91$) configurations, $p = 0.781$. Dynamic spectral

Figure 8. Post hoc comparison for the perception of engulfment for the four spatialization techniques.



subband decorrelation was significantly different between horizontal-only ($\bar{x} = 3.1$, $s = 1.09$) and elevated-only ($\bar{x} = 3.8$, $s = 1.09$) configurations $p < 0.05$, but it was not significant between horizontal-only ($\bar{x} = 3.1$, $s = 1.09$) and 3-D ($\bar{x} = 3.6$, $s = 1.00$) configurations, $p = 0.55$.

Spectral splitting was not significantly different between horizontal-only ($\bar{x} = 2.8$, $s = 0.85$) and elevated-only ($\bar{x} = 3.9$, $s = 1.19$) configurations, $p = 0.380$, but was found to be significantly different between the horizontal-only ($\bar{x} = 2.8$, $s = 0.85$) and 3-D ($\bar{x} = 3.5$, $s = 0.95$) configurations, $p < 0.05$. Timbre spatialization was significantly different between horizontal-only ($\bar{x} = 3.4$, $s = 0.95$) and elevated-only ($\bar{x} = 2.8$, $s = 0.917$) configurations, $p < 0.05$ and was not significantly different between the horizontal-only ($\bar{x} = 3.4$, $s = 0.95$) and 3-D ($\bar{x} = 2.9$, $s = 1.19$) configurations, $p = 0.96$.

Discussion

Based on the results for the main effect for loudspeakers, the use of horizontal-only loudspeaker configurations contributes to the perception of envelopment, and so indicates that Null Hypothesis 1 is rejected for envelopment.

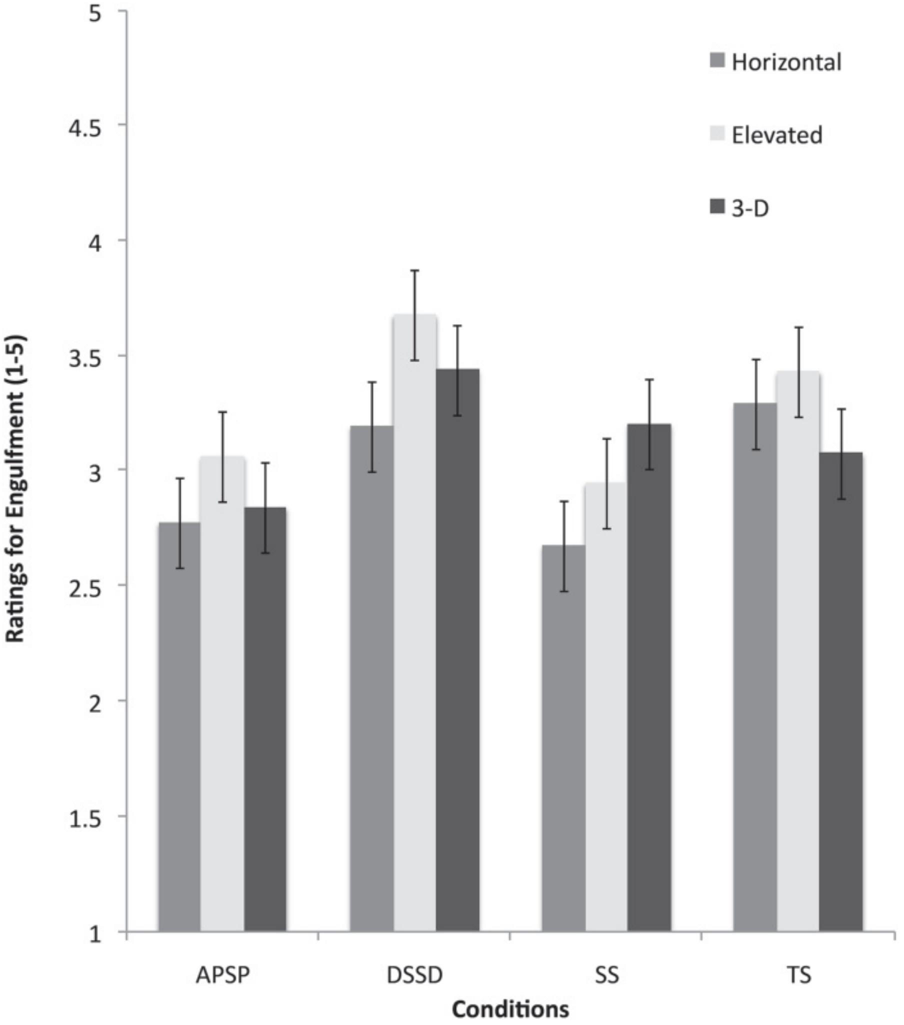
Based on the results for the main effect for techniques, the use of DSSD contributes to the perception of envelopment significantly more than the other techniques evaluated. This suggests that listeners perceive DSSD as the most enveloping technique for the perception of envelopment, and so Null Hypothesis 2 is rejected for envelopment.

The data for the main interaction effect for techniques and loudspeakers indicates no significant dependency between these two variables for the perception of envelopment. This suggests that the loudspeaker configurations did not affect listeners' ratings for the spatialization techniques tested. Apart from TS, which was shown to be significantly different between the horizontal-only and elevated-only loudspeaker configurations, no other technique showed a significant result between loudspeaker configurations. Thus, Null Hypothesis 3 was not rejected for envelopment. Interestingly, all techniques received the highest mean ratings in the horizontal-only configuration and lowest mean ratings in either the elevated-only or the 3-D configurations for envelopment.

As for the perception of engulfment, the data for the main effect of loudspeaker configurations suggests no significant difference between loudspeakers. This indicates that the loudspeaker configuration had no effect on the listener's perception of engulfment, rejecting Null Hypothesis 1. Nevertheless, the highest mean rating for engulfment was for the 3-D configuration, followed by the elevated-only and horizontal-only configurations, respectively.

The results for the main effect of techniques suggest that the use of DSSD contributes to the perception of engulfment significantly more the other techniques. This suggests that listeners perceive DSSD as the most engulfing technique, and Null Hypothesis 2 is rejected for engulfment.

Figure 9. Participants rating for the perception of engulfment for the spatialization techniques and loudspeaker configurations.



The data for the main interaction effect between techniques and loudspeakers suggest a significant dependency between these two variables for the perception of engulfment. This suggests that the perception of techniques for levels of engulfment is dependent on loudspeaker configuration. Listeners perceived DSSD to be significantly more engulfing when perceived in the elevated-only configuration than in the horizontal-only configuration. For spectral splitting, the listeners perceived the technique to be significantly more engulfing in the 3-D configuration than in the horizontal-only configuration.

Conclusion

Based on these results, (1) DSSD was rated highest for levels of envelopment and engulfment, (2) there is a strong correlation between the perception of envelopment and the use of the horizontal-only loudspeaker configurations, and (3) the perception of engulfment in elevated-only configurations is reliant on the spatialization technique used. The only technique that demonstrated a significant difference for engulfment between horizontal-only and elevated-only configurations was DSSD. In addition, spectral splitting was significantly different for the perception of engulfment between the horizontal-only and 3-D configurations, enforcing the technique's aim as a 3-D spatialization approach. Because DSSD was perceived by listeners to be the most enveloping and engulfing technique, we chose to develop upon it further, with the objective of increasing its perceived level of envelopment and engulfment. We have since used this technique to compose 3-D electroacoustic works.

Acknowledgments

This research would not have been possible without the support of students and staff at the Digital Media and Arts Research Centre, Department of Computer Science and Information Systems, University of Limerick, Ireland. Special thanks to Kerry Hagan, Darragh Pigott, and Annette McElligott.

References

Adair, S., M. Alcorn, and C. Corrigan. 2008. "A Study into the Perception of Envelopment in Electroacoustic Music." In *Proceedings of International Computer Music Conference*, pp. 342–345.

Barreiro, D. 2010. "Considerations on the Handling of Space in Multichannel Electroacoustic Works." *Organised Sound* 15(01):290–296.

Barron, M. 1971. "The Subjective Effects of First Reflections in Concert Halls: The Need for Lateral Reflections." *Journal of Sound and Vibration* 15(4):475–494.

Barron, M., and A. Marshall. 1981. "Spatial Impression Due to Early Lateral Reflections in Concert Halls: The Derivation of a Physical Measure." *Journal of Sound and Vibration* 77(2):211–232.

Basque, N., and A. Watson. 2004. "Présentation des œuvres électroacoustiques sur des systèmes multi haut-parleurs." *eContact!* 7(4). Available online at econtact.ca/7_4/watson-basque_dhomont.html. Accessed November 2016.

Beranek, L. 1992. "Concert Hall Acoustics: 1992." *Journal of the Acoustical Society of America* 92(1):1–39.

Berg, J., and F. Rumsey. 1999. "Identification of Perceived Spatial Attributes of Recordings by Repertory Grid Technique and Other Methods." Paper presented at the Audio Engineering Society Convention, 8–11 May, Munich, Germany.

Berg, J., and F. Rumsey. 2001. "Verification and Correlation of Attributes Used for Describing the Spatial Quality of Reproduced Sound." In *Proceedings of Audio Engineering Society Conference: Surround Sound: Techniques, Technology and Perception*, pp. 234–251.

Berg, J., and F. Rumsey. 2002. "Validity of Selected Spatial Attributes in the Evaluation of 5-Channel Microphone Techniques." Paper presented at the Audio Engineering Society Convention, 8–13 May, Munich, Germany.

Bradley, J. 1994. "Comparison of Concert Hall Measurements of Spatial Impression." *Journal of the Acoustical Society of America* 96(6):3525–3535.

Bradley, J., and G. Soulodre. 1995. "Objective Measures of Listener Envelopment." *Journal of the Acoustical Society of America* 98(5):2590–2597.

Bridger, M. 1989. "An Approach to the Analysis of Electro-Acoustic Music Derived from Empirical Investigation and Critical Methodologies of Other Disciplines." *Contemporary Music Review* 3(1):145–160.

Cheng, C., and G. Wakefield. 2001. "Moving Sound Source Synthesis for Binaural Electroacoustic Music Using Interpolated Head-Related Transfer Functions (HRTFs)." *Computer Music Journal* 25(4):57–80.

Chowning, J. 1971. "The Simulation of Moving Sound Sources." *Journal of the Audio Engineering Society* 19(1):2–6.

Davis, T., and P. Rebelo. 2005. "Hearing Emergence: Towards Sound Based Self- Organization." In *Proceedings of International Computer Music Conference*, pp. 463–466.

Delalande, F. 1989. "La Terrasse des audiences du clair de lune: Essai d'analyse esthésique." *Analyse Musicale* 15:75–85.

Deliège, I. 1989. "A Perceptual Approach to Contemporary Musical Forms." *Contemporary Music Review* 4(1):213–230.

Desantos, S., C. Roads, and F. Bayle. 1997. "Acoustmatic Morphology: An Interview with François Bayle." *Computer Music Journal* 21(3):11–19.

Emmerson, S. 2007. "Diffusion–Projection: The Grain of the Loudspeaker." In *Living Electronic Music.* Hampshire, UK: Ashgate, pp. 143–170.

Griesinger, D. 1999. "Objective Measures of Spaciousness and Envelopment." In *Proceedings of Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction*. Paper 16-003 (pages unnumbered).

Guastavino, C., and B. Katz. 2004. "Perceptual Evaluation of Multi-Dimensional Spatial Audio Reproduction." *Journal of the Acoustical Society of America* 116(2):1105–1115.

Hamasaki, K., K. Hiyama, and R. Okumura. 2005. "The 22.2 Multichannel Sound System and Its Application." In *Proceedings of Audio Engineering Society 118th Convention*. Paper 6406 (pages unnumbered).

Harrison, J. 1998. "Sound, Space, Sculpture: Some Thoughts on the 'What', 'How' and 'Why' of Sound Diffusion." *Organised Sound* 3(2):117–127.

Harrison, J. 1999. "Imaginary Space: Spaces in the Imagination." Keynote Address presented at the Australasian Computer Music Conference, 7–10 July, Victoria University of Wellington, New Zealand. Reprinted in *eContact!* 3(2). Available online at econtact.ca/3_2/ACMConference.htm. Accessed December 2016.

Holman, T. 2008. *Surround Sound: Up and Running.* 2nd ed. Oxford: Focal.

Kahle, E. 1995. "Validation of an Objective Model for Characterizing the Acoustic Quality of a Set of Concert Hall and Opera Houses." PhD dissertation, Université du Maine, Paris and Laboratoire d'Acoustique du Maine, IRCAM.

Kendall, G. 1995. "The Decorrelation of Audio Signals and Its Impact on Spatial Imagery." *Computer Music Journal* 19(4):71–87.

Kendall, G., and A. Cabrera. 2011. "Why Things Don't Work: What You Need to Know About Spatial Audio." In *Proceedings of the International Computer Music Conference*, pp. 37–40.

Kim, S., Y. Lee, and V. Pulkki. 2010. "New 10.2-Channel Vertical Surround System (10.2-VSS): Comparison Study of Perceived Audio Quality in Various Multichannel Sound Systems with Height Loudspeakers." In *Proceedings of Audio Engineering Society 129th Convention*. Paper 8296 (pages unnumbered).

Kim-Boyle, D. 2005. "Sound Specialization with Particle Systems." In *Proceedings of International Conference on Digital Audio Effects*, pp. 65–68.

Kim-Boyle, D. 2006. "Spectral and Granular Spatialization with Boids." In *Proceedings of the International Computer Music Conference*, pp. 139–142.

Landy, L. 1994. "The 'Something to Hold on to Factor' in Timbral Composition." *Contemporary Music Review* 10(2):49–60.

Lee, H., and C. Gribben. 2014. "Effect of Vertical Microphone Layer Spacing for a 3D Microphone Array." *Journal of the Audio Engineering Society* 62(12):870–884.

Lynch, H., and R. Sazdov. 2011a. "An Ecologically Valid Experiment for the Comparison of Established Spatial Techniques." In *Proceedings of International Computer Music Conference*, pp. 130–134.

Lynch, H., and R. Sazdov. 2011b. "An Investigation into the Compositional Techniques Utilized for the Three-Dimensional Spatialization of Electroacoustic Music." In *Proceedings of Electroacoustic Music Studies Conference, Sforzando*! (EMS 2011). Available online at ems-network.org/spip.php?article328. Accessed November 2016.

McCartney, A. 2000. "Sounding Places with Hildegard Westerkamp." PhD dissertation, Electronic Music Foundation Institute.

Morimoto, M. 1997. "The Role of Rear Loudspeakers in Spatial Impression." Paper presented at the Audio Engineering Society Convention, 26–29 September, New York City.

Morimoto, M., K. Iida, and L. Sakagami. 2001. "The Role of Reflections from behind the Listener in Spatial Impression." *Applied Acoustics* 62(2):109–124.

Nakayama, T., et al. 1971. "Subjective Assessment of Multichannel Reproduction." *Journal of the Audio Engineering Society* 19(9):744–751.

Normandeau, R. 2009. "Timbre Spatialization: The Medium Is the Space." *Organised Sound* 14(03):277–285.

Peters, N., G. Marentakis, and S. McAdams. 2011. "Current Technologies and Compositional Practices for Spatialization: A Qualitative and Quantitative Analysis." *Computer Music Journal* 35(1):10–27.

Potard, G., and I. Burnett. 2004. "Decorrelation Techniques for the Rendering of Apparent Sound Source Width in 3D Audio Displays." In *Proceedings of the International Conference on Digital Audio Effects*. Available online at ant-s4.unibw-hamburg.de/dafx/paper-archive/2004/P_280.PDF. Accessed November 2016.

Rolfe, C. 1999. "A Practical Guide to Diffusion/Projection." Available online at econtact.ca/2_4/pracdiff.htm. Accessed November 2016.

Ronan, M., D. Piggott, and R. Sazdov. 2012. "Configuring SpADE: Practical Considerations Influencing the Design of a Three-Dimensional Multi-Channel Listening Environment." Paper presented at the Irish Sound, Science and Technology Convocation, 1–2 August, Cork, Ireland.

Rumsey, F. 1998. "Subjective Assessment of the Spatial Attributes of Reproduced Sound." In *Proceedings of Audio Engineering Society Conference: Audio, Acoustics and Small Space*, pp. 122–135.

Rumsey, F. 2002. "Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm." *Journal of the Audio Engineering Society* 50(9):651–666.

Sazdov, R. 2011a. "An Ecologically Valid Experiment for the Perceptual Study of Multi-Channel Electroacoustic Music." In A. Vande Gorne, ed. *L'Espace du son III*. Brussels: Musiques & Recherches, pp. 71–90.

Sazdov, R. 2011b. "The Effect of Elevated Loudspeakers on the Perception of Engulfment, and the Effect of Horizontal Loudspeakers on the Perception of Envelopment." Paper presented at the International Conference on Spatial Audio, 10–13 November, Detmold, Germany.

Sazdov, R. 2015. "Envelopment vs. Engulfment: Multidimensional Scaling on the Effect of Spectral Content and Spatial Dimension within a Three-Dimensional Loudspeaker Setup." Paper presented at the International Conference on Spatial Audio, 17–20 September, Graz, Austria.

Sazdov, R., G. Paine, and K. Stevens. 2007. "Perceptual Investigation into Envelopment, Spatial Clarity and Engulfment in 3D Reproduced Multi-Channel Loud-Speaker Configurations." Paper presented at the Audio Engineering Society Conference, 25–27 June, London.

Schroeder, M., D. Gottlob, and K. Siebrasse. 1974. "Comparative Study of European Concert Halls: Correlation of Subjective Preference with Geometric and Acoustic Parameters." *Journal of the Acoustical Society of America* 56(1):1195–1201.

Silzle, A., G. Sunish, T. Bachmann. 2011. "Experimental Setups for 3D Audio Listening Tests." Paper presented at the International Conference on Spatial Audio, 10–13 November, Detmold, Germany.

Smalley, D. 1996. "The Listening Imagination: Listening in the Electroacoustic Era." *Contemporary Music Review* 13(2):77–107.

Soulodre, G., M. Lavoie, and S. Norcross. 2002. "Investigation of Listener Envelopment in Multichannel Surround Systems." Paper presented at the Audio Engineering Society Convention, 5–8 October, Los Angeles, California.

Stockhausen, K. 1971. "Osaka-Projekt." In *Texte zur Musik 1963–1970*. Cologne: DuMont, pp. 153–187.

Torchia, R., and C. Lippe. 2004. "Techniques for Multi-Channel Real-Time Spatial Distribution Using Frequency-Domain Processing." In *Proceedings of Conference on New Interfaces for Musical Expression*, pp. 116–119.

Weale, R. 2005. "The Intention/Reception Project: Investigating the Relationship between Composer Intention and Listener Response in Electroacoustic Compositions." PhD dissertation, De Montfort University.

Wilson, S. 2008. "Spatial Swarm Granulation." Paper presented at the International Computer Music Conference, 24–29 August, Belfast, UK.

Wilson, S., and J. Harrison. 2010. "Rethinking the BEAST: Recent Developments in Multichannel Composition at Birmingham ElectroAcoustic Sound Theatre." *Organised Sound* 15(3):239–250.

Wyatt, S. 1999. "Investigative Studies on Sound Diffusion/Projection." Available online at ems.music.illinois.edu/ems/articles/sound_projection. Accessed November 2016.

Zacharov, N., and K. Koivuniemi. 2001. "Unravelling the Perception of Spatial Sound Reproduction: Language Development, Verbal Protocol Analysis and Listener Training." Paper presented at the Audio Engineering Society Convention, 30 November–3 December, New York City. Available online at www.aes.org/e-lib/browse.cfm?elib=9815 (subscription required). Accessed December 2016.