CRANIAL TRANSITIONS FOR SOPRANO SAXOPHONE AND ELECTRONIC PROCESSING

Jonas Braasch

School of Architecture
Rensselaer Polytechnic Institute
110 8th St, Troy, NY 12180
braasj@rpi.edu

ABSTRACT

Cranial Transitions was written for soprano saxophone and electronic processing to explore the relationship between internally and externally perceived auditory events. The piece, which was recorded for reproduction with headphones, is based on the personal observation that too often our internal representation of musical ideas seems to be more ambiguous and abstract than we share with others. Thus, the external auditory events which symbolize the outside representation of our musical life follow more traditional techniques of arpeggios and melodies, while the internal representation follows more abstract ideas. The internal world is manipulated with electronic processing using a specially designed program, the *Intra-Cranial Spatializer*. In contrast, the outside world captures the acoustic side of the soprano saxophone. In this case, electronic processing is only used to create a virtual acoustic environment.

[Keywords: Binaural Technology, Spatial Music]

1. INTRODUCTION

By definition, spatial parameters play a central role in the genre of Spatial Music. Usually, either musical instruments are positioned at different locations or electronic sound sources are projected through an array of spatially placed loudspeakers. The combination of both techniques within one piece is also common. In both cases, the auditory events are perceived outside of the head (extra-cranial event), due to the external location of the physical sound sources (acoustical instruments or loudspeakers). In other cases, however, a sound source can be perceived within the head (the so-called *inside-the-head locatedness*) to form an intra-cranial event. This effect can be observed when a monophonic sound source (e.g., AM radio) is rendered through a pair of headphones in mono. The phenomenon that two identical (diotic)[1] headphone signals lead to intra-cranial events in the center of the head has been described in psychoacoustic literature from early on (e.g., [22]). Inside-the-head located auditory events plays a central role in the design of psychoacoustic experiments (e.g., lateralization experiments), but is unwanted in music playback. While the desire to achieve an extra-cranial music experience is understandable for traditional music, our intra-cranial space is certainly worth exploring from an avant-garde standpoint—in particular in relationship to the extra-cranial space. The concept of projecting sounds both internally and externally goes beyond Pauline Oliveros' piece *Arctic Air* (1992) where the performers where instructed to "play what is sounding just inside the ear" vs. to "play what is sounding just outside the ear." It is noteworthy that an analogon to an intra-cranial sound projection does not exist in vision—apart from drug induced cases. Also Stelarc's *Stomach Sculpture* does not lead to a perceived internal visual image, but only to one that is felt.

[1]The term *diotic* refers to a stimulation of both ears with identical signals. The term *dichotic* is used for a presentation at both ears with two different signals, while *monotic* refers to the acoustic stimulation of only one ear.

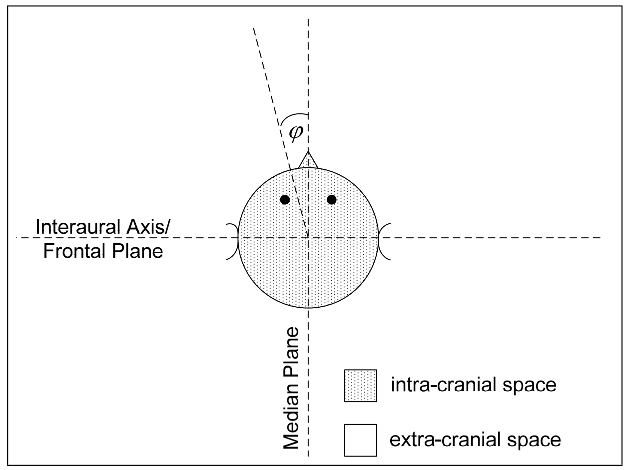

Figure 1: Head-related coordinate system (Top view).

The piece *Cranial Transitions* is an exploration of the intra- and extra-cranial world and the comparison of both spaces. This paper describes the main concepts and underlying techniques that were used to navigate through both worlds. After a brief introduction of the *head-related coordinate system,* which is used throughout this article, the psychoacoustical principles that were utilized in *Cranial Transitions* are outlined, followed by the description of the actual piece.

2. THE HEAD-RELATED COORDINATE SYSTEM

The origin of the head-related coordinate system is located on the *interaural axis*—which intersects the upper margins of the entrances to left and right ear canals—halfway between the entrances to the ear canals (see Figs. 1 and 2). The *horizontal plane* is defined by the interaural axis and the lower margins of the eye sockets, while the *frontal plane* lies orthogonal on the horizontal plane intersecting the interaural axis. The *median plane* is orthogonal to

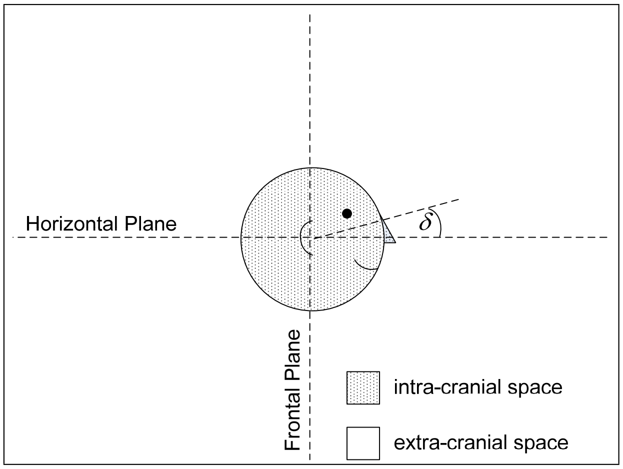

Figure 2: Head-related coordinate system (Side view).

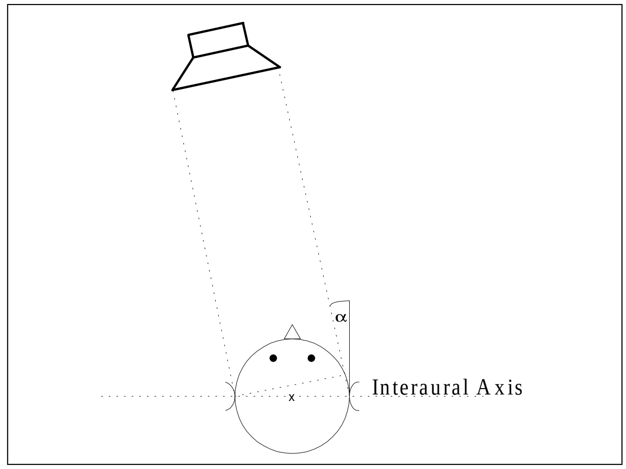

Figure 3: Simple geometric model to estimate interaural time differences [17].

the horizontal plane, as well as to the frontal plane and therefore, so to say, cuts the head in two symmetrical halves.[2] The position of a sound source is described using the polar coordinates: azimuth φ, elevation δ, and distance d. If δ is zero and d is positive, the sound source moves anti-clockwise through the horizontal plane with increasing φ.[3] At φ=0° and δ=0°, the sound source is directly in front of the listener, intersecting the horizontal and median plane. If φ is zero and d is positive, the sound source moves in front of the listener with increasing δ, upwards along the median plane and downwards behind the listener. In Figures 1 and 2, the intra-cranial areas are shown hatched, while the extra-cranial areas have a white background.

3. PSYCHOACOUSTIC BACKGROUND

For many centuries, scientists have been trying to understand how the human auditory system is able to localize sound sources in space. As early as 1882, S. P. Thompson [32] wrote a review of different theories with respect to this topic, where he listed up what can be regarded the fundamental laws of binaural hearing up to the present day. According to Thompson, the sensitivity of the auditory system to interaural phase differences had been found in 1877 by Bell, W. Thompson and himself [31], independently from each other. Their findings led to a theory according to which the arrival times of the sound wave emitted from a single source are usually not exactly the same at the left and right eardrums—due to the different path-lengths to both ears. This arrival-time difference between the left and right ear is called *interaural time difference* (ITD). With a simple geometric model [17], see Fig. 3, it can be shown that the maximal ITD is measured when the sound wave arrives from the side along the axis which intersects both eardrums. In this case, the ITD can be estimated as the distance between the eardrums, $\approx 18\,\mathrm{cm}$, divided by the speed of sound, $\approx 340\,\mathrm{m/s}$, to a value of $529\,\mu\mathrm{s}$.[4] However, larger ITDs than those are observed in nature. Because of shadowing effects of the head, the measured ITDs can be, depending on the head size, as large as $800\,\mu\mathrm{s}$. Taking the model of Hornbostel and Wertheimer [17], ITDs of the same magnitude form hyperbolas in the horizontal plane, and at greater distances the shell of a cone in three-dimensional space is apparent, the so-called *cones of confusion*. Hence, there exist multiple positions with identical ITDs—yet despite of this, the ITDs are still very reliable cues to determine the left–right lateralization of a sound source. A model which estimates the ITDs on the basis of the wave travelling around a sphere has been proposed in [34], which is still a good prediction in the high frequency range. Later this model was modified to predict the ITDs for all frequencies throughout the human hearing range [21].

The existence of the head between both ears does not only determine the detour the traveling sound wave has to follow, but also causes attenuation of the sound wave at the contralateral eardrum, which leads to *interaural level differences* (ILDs)[5] between both ear signals. Already in the end of the 19^{th} century, a geometric model was established to estimate ILDs for various sound-source positions [30]. In contrast to the ITDs, the ILDs are strongly frequency dependent. In the low frequency range, the human head is small in comparison to the wave length and, therefore, diffraction has only a minor effect on the sound wave. In the high frequency range, however, the wave length is short as compared to the dimensions of the head, and much larger ILDs than in the low frequency range can be observed. In this frequency region, the ILDs are not only determined by the shape of the head, but are also greatly influenced by the shape of the outer ears.

The frequency dependence of the ILDs led to the idea of the duplex theory which claims that ITDs are the dominant cue in the low frequency range, while ILDs are more important than ITDs at high frequencies. Lord Rayleigh [27] showed both theoretically and in a psychoacoustic experiment that the head is very effective in attenuating the sound at the contra-lateral ear for high, but not for low frequencies. For this reason, ILDs are considered too small in the low frequency range to provide a reliable localization cue. For ITDs, it was concluded that the unequivocal relationship be-

[2] It should be noted that the head is not perfectly symmetrical. For this reason slight interaural time and level differences are also measured for sound sources in the median plane.

[3] Note that in the field of audio engineering, the azimuth is denoted clockwise.

[4] Hornbostel and Wertheimer estimated the distance between the two eardrums to 21 cm. This value, however, is too large, and nowadays it is common to use 18 cm instead.

[5] Note that interaural level differences are frequently referred to as interaural intensity differences, IIDs.

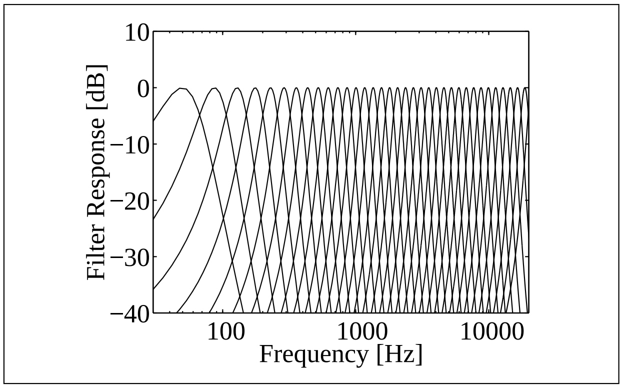

Figure 4: Magnitude responses of an auditory filter bank. A gammatone-filter-bank implementation according to [25] is shown with 36 bands.

tween the phase difference between both ear signals and auditory lateralization vanishes for high frequencies, where the path-length difference between both ears exceeds the wave length of signals stemming from a sideway sound source. Indeed, it was later shown for sinusoidal test signals [24] that our auditory system is not able to resolve the fine structure of signal frequencies above approximately 1.5 kHz. Nowadays, however, the duplex theory is not seen in such a strict way anymore as it was originally proposed. It is now being realized that ITDs are important in the high frequency range as well, as they can be evaluated through envelope fluctuations [10].

The question of how ITDs and ILDs are combined in the auditory system to estimate the position of the sound source has been only partly answered so far. For a long period of time, it was assumed that ITDs and ILDs are evaluated separately in the auditory system, namely, ITDs in the medial superior olive, MSO, as was first shown for dogs [14], and ILDs in the lateral superior olive, LSO. Yet, recent neurophysiological findings have shown that the ITDs in the envelopes of modulated sounds, and even in low-frequency carriers, are also processed in the LSO [18], [19]. In addition, the occurrence of the so-called *time-intensity-trading effect* revealed a high complexity of the processing in the auditory system already half a century ago [10], [15]. The trading effect describes the phenomenon that the auditory event often evolves midway between the positions with ITD and ILD cues leading into opposite direction and such compensate for each other. It has been suggested [18] that the combined sensitivity of single neurons in the LSO to ITDs and ILDs offers an easy explanation for the time-intensity trading effect. However, the auditory event of the listeners become spatially diffuse or they even split up into more than one auditory event when the ITDs and ILDs differ too much from the “natural” combinations of ITDs and ILDs as observed in free-field listening (e.g., [13]).

For a more detailed understanding of how the interaural cues are analyzed across frequency it is important to know that the auditory system analyzes sounds in overlapping frequency bands [12] (see Fig. 4). The widths of these frequency bands always correspond to a constant distance of approximately 2 mm on the basilar membrane. The relationship of frequency and distance on the basilar membrane between the place of maximum deflection and the helicotrema is approximately logarithmic. Therefore, the frequency bands become broader with the frequency. The frequency

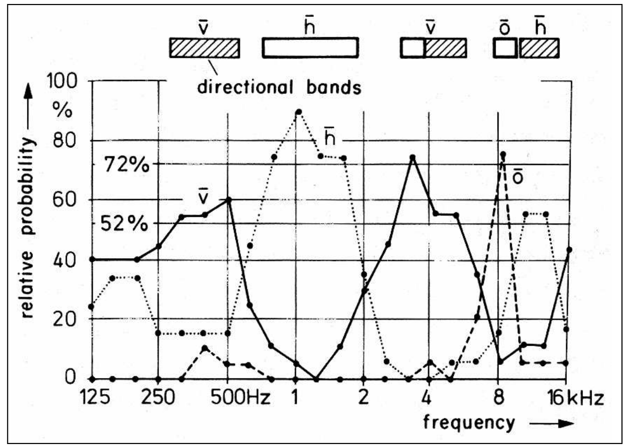

Figure 5: Relative probability of listeners’ direction judgements (front, v; above, o; or rear, h) in response to 1/3-octave noise bursts with different center frequencies. The top of the graph shows those directions for which the listeners reported this direction in more than 50% of the cases. The data is averaged over 10 listeners (from Blauert [2], [5]).

selective analysis is especially important for ILDs, because of their strong frequency dependence. By evaluating the ILDs and ITDs across several frequency bands, an unequivocal sound-source position can be easily determined for most directions. The preservation of the “natural” combinations of ILDs and ITDs across frequency are also important for the external perception of an acoustic event. Hartmann and Wittenberg [16] found that for ILDs the whole frequency range is relevant to this process, while for ITDs only frequencies up to 1 kHz are important. Also wall reflections are commonly known to support the externalization of auditory events, and the energy ratio between the direct sound source and its reflections is a strong cue in distance perception [35].

The median plane plays a special role in binaural psychoacoustics, as the interaural cues are very small here and cannot reliably be used to resolve positions. In this case, other cues become important. Blauert presented narrowband noise bursts to 10 listeners from three different loudspeaker locations (front, rear or above) in an anechoic room [2]. The listeners were asked to report the perceived direction of incident. Interestingly, the responses did not correlate much with the actual loudspeaker positions, but were rather a function of the center frequency of the stimuli (Fig. 5). Blauert concluded that for different elevation angles, the spectrum of the ear signals is characteristically boosted or attenuated in different frequency bands, due to diffraction and scattering at the head and outer ears. Measurements revealed that the measured boosted bands correlated well with the perceived directions for narrow-band stimuli at these center frequencies (see Fig. 6 in comparison to Fig. 5).

It is assumed that the auditory system performs a spectral analysis to determine the position within the median plane. This theory is based on the finding that, for sinusoidal and other narrow-band sound sources, the positions of the auditory events are formed in positions for which the signals to the ears show local maxima for broad-band sound sources, so called *Directional Bands* or *Blauert Bands*. Spectral cues are called *monaural cues*, since one ear only is sufficient to resolve them.

Already in 1948, Licklider [22] described an experiment in

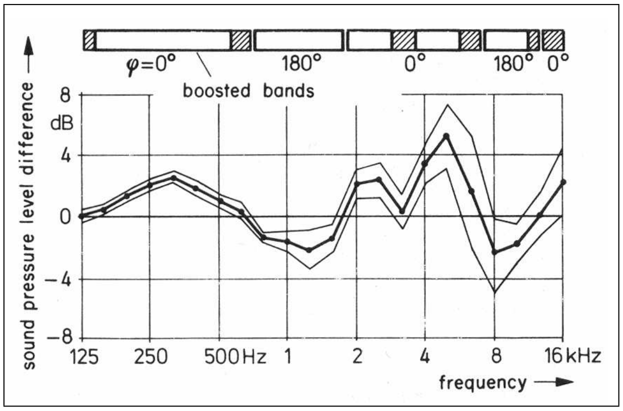

Figure 6: Frequency-dependent sound-pressure level differences between frontal and rear sounds as measured at the ear-canal entrances of human subjects. The boosted bands are shown on top of the graph (white areas: 95% confidence; shaded areas: most probable cases (from Blauert [2], [5]).

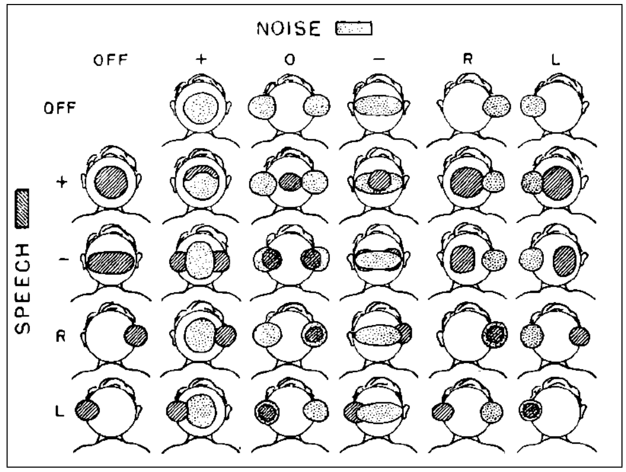

Figure 7: Intra-cranial events for a speech and a noise signal under various binaural conditions: '+', in phase; '−', out of phase; '0', random phase; 'R', right channel; 'L', left channel. Each image displays the rear view of the listener's head (from Licklider [22]).

which a noise signal and a speech signal were perceived at different intra-cranial positions, depending on the interaural phase relationships between both signals (Fig. 7). Licklider generated the stimuli for his experiment with the analog circuit depicted in Fig. 8. The figure shows the circuit to create the interaural noise signals. The circuit for the speech signals is very similar. The boxes labeled with C, D, and E represent three independent noise generators. The outputs x and y were fed to the left and right headphone channels. The circuit provided the option to phase invert the right channel for noise generator C. Although, Licklider investigated primarily the perceptual separation of a speech and a noise signal for various binaural conditions, his study also demonstrates cases for multiple intra-cranial auditory events. Later, Chernyak and Dubrovsky [9] and others further investigated how the degree of coherence k between both ear signals affected the perceived width of the intra-cranial pattern using a device similar to Licklider's (Fig. 9). The perceived auditory event widens with decreasing correlation until it splits into two separate auditory events (Fig. 9, bottom right). Other experimenters showed that a signal presented via headphones can be shifted left and right along the interaural axis if ILDs or ITDs are applied (e.g., [29],[33]). Blauert and Lindemann [4] suggested that *spaciousness*, the spatial extent of the auditory event, is increased by rapid ILD and ITD fluctuations of the stimulus.

4. CREATION OF EXTRA-CRANIAL EVENTS

The dummy-head recording technique was chosen to record the saxophone takes. Since the dummy-head dimensions are similar to those of a human head and upper torso, the generated binaural cues at ear signal level (entrance of the ear canals) are very similar to those found for humans. The recording technique typically leads to externally perceived events, if the dummy-head outputs are monitored via headphones. Although, individual differences in head and pinnae dimensions of a human listener compared to the dimensions of the recording head can lead to perceptual artifacts such as front/back confusions, the dummy-head recording technique surpasses any other microphone technique for headphone reproduction that can be captured with commercial media such as the Compact Disk. The use of headphones for playback was a requirement for this piece, since loudspeaker based techniques are hardly suitable to produce intra-cranial events. A Head Acoustics HMS II.1 dummy head was used for this project. In contrast to most commercial binaural recordings, the takes for *Cranial Transitions* were obtained in an acoustically treated room with hardly any wall reflections. The latter were added later using the mirror-image technique described in the next section. The advantage of this procedure is two fold:

Firstly, recordings between intra- and extra-cranial projections could be obtained seamlessly, which was important for the musical context. Otherwise, it would have been necessary to record the parts for the internal events separately, since a reverberation free environment was desired to avoid accidental externalization of the auditory events.

Secondly, the procedure allowed to adjust the amount of reverberation and choice of room type in the post-production phase. In particular, it was possible to simulate greater distances than the actual recording distance by adding more reverberation.

4.1. Impulse Response Generation

In order to create the virtual reverberant environment, a rectangular room with variable dimensions was simulated using the mirror-image method [1]. To simulate early reflections, the first 200 wall reflections were modeled (up to third order reflections, plus a few fourth order reflections). Each reflected mirror source was filtered with the head-related transfer functions (HRTFs)[6], measured at the closest available angle. The dummy head recording provided the HRTFs for the direct sound source. The HRTFs for the room impulse response were measured on the author in the anechoic chamber of the Institute of Communication Acoustics at the Ruhr University Bochum, Germany (see [7] for a description of the measurement procedure). The sound source was presented at

[6]The head-related transfer functions (HRTFs) describe the transfer function from a position in free-field space to the left and right entrance of the ear canal.

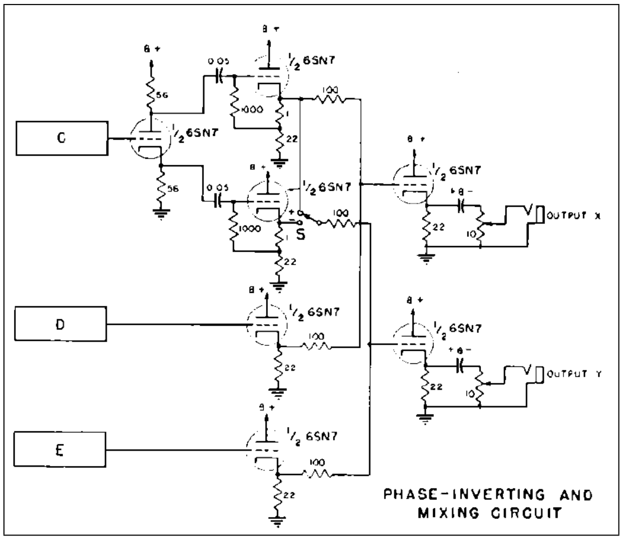

Figure 8: Circuit to generate a two-channel noise signal with adjustable interaural phase relationships (from Licklider [22]).

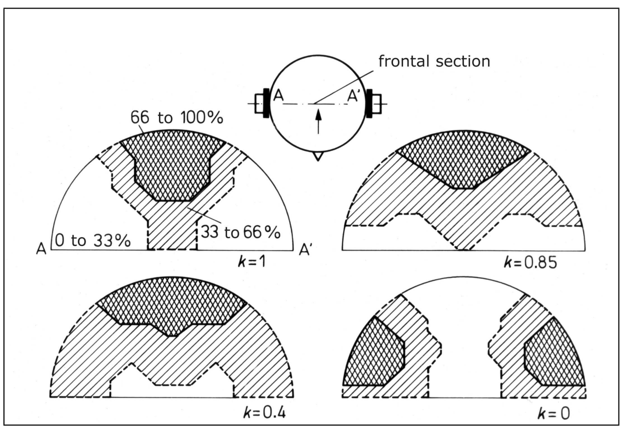

Figure 9: Intra-cranial auditory events for a dichotic noise signal for different degrees of coherence (from Blauert [5] after Chernyak and Dubrovsky [9]).

a distance of $5\,\mathrm{m}$ from the virtual listener, who was placed in the center of the virtual room, at the same height as his ears ($1.30\,\mathrm{m}$). The frequency dependent absorption coefficients of the walls and the floor are taken from measurements described in the Deutsche Industrie Norm (DIN, German Industrial Norm) [11].

The late reverberation pattern was generated using a stereo sample of white noise. The signal was octave-band filtered with center frequencies at $125\,\mathrm{Hz}$ (low-pass filter), $250\,\mathrm{Hz}$, $500\,\mathrm{Hz}$, $1\,\mathrm{kHz}$, $2\,\mathrm{kHz}$, $4\,\mathrm{kHz}$, $8\,\mathrm{kHz}$, $16\,\mathrm{kHz}$ (high-pass filter). An exponentially decaying envelope was applied to each frequency band. The reverberation time for each envelope was determined using Sabine's law:

$$T_r = 0.163\frac{V}{S\hat{\alpha}} \tag{1}$$

with the reverberation time T_r, the volume of the space V, the surface area S, and the average absorption coefficient $\hat{\alpha}$. The maximum amplitude of the envelope is set to the amplitude of the direct

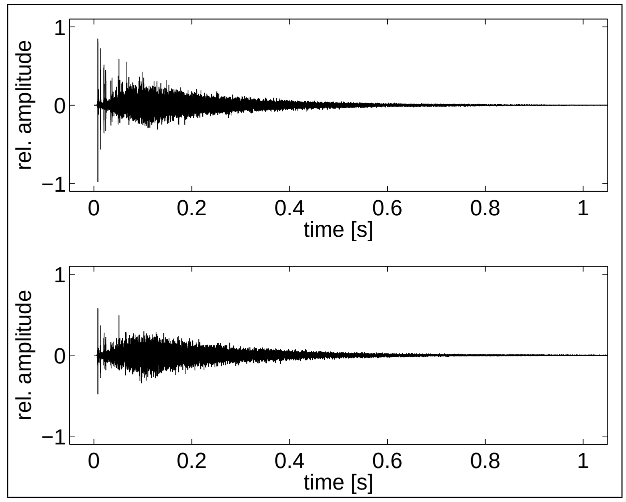

Figure 10: Artificial room impulse response (RT=$2.0\,\mathrm{s}$) generated using the mirror-image technique and an HRTF catalog: left channel, top graph; right channel, bottom graph.

sound. Next, the portion before the arrival time of the earliest reflection plus an interval of $2\,\mathrm{ms}$, t_1, is set to zero. A linear slope from zero at t_1 to one at t_{200}, the arrival time of the 200$^{\text{th}}$ reflection, is applied to scale the decay curve during this time interval. Afterward, the reverberation tail decreases exponentially (weighting factor of one). An example of the artificially created room impulse response is shown in Fig. 10.

The impulse response was convolved with the signal using the SIR VST plugin (ver. 1.011) [20]. Since the input signal was taken from the dummy-head recording, the room impulse response had to be modified to avoid double HRTF filtering (spectral filtering through the dummy head and HRTF filtering during the room impulse calculation). An equalization filter was applied to the impulse response to compensate for the frequency alteration (spectral de-emphasis in the high frequency range) of the dummy head's pinnae. This was achieved using a parametric filter to enhance the region around $11\,\mathrm{kHz}$ by $9\,\mathrm{dB}$ (Q=1.8). A more accurate procedure would have been to apply inverse filtering using the dummy head's own HRTFs. Considering the circumstance that dynamic source movements were part of the dummy-head recording, the inverse filtering with the HRTFs would not have been feasible. In particular, characteristic notches in the HRTFs, which vary in frequency for different sound-source positions, could have led to unwanted peaks in the equalized signal, if the notches of the HRTF equalization filter had not matched exactly the characteristic frequency notches that were induced by the dummy head during the recording process.

5. CREATION OF INTRA-CRANIAL EVENTS

As stated earlier, one of the main concerns in *Cranial Transitions* was to ensure that the intra- and extra-cranial events were perceived as such. To avoid that the intra-cranial events are perceived externally, the related parts of the dummy head recording were processed according to the following guidelines:

- Natural combinations of ITDs and ILDs as found in the HRTFs were not considered, since sound sources tend to be perceived externally, if the localization parameters follow

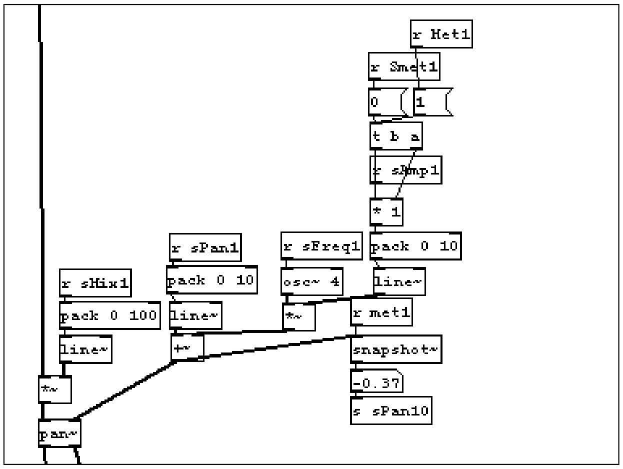

Figure 11: Pure Data patch to generate intra-cranial movements.

natural patterns. Therefore, ITDs and ILDs were removed by transforming the binaural signal to mono by deleting the left channel.

- The sound source was then artificially lateralized based on interaural level differences only.
- The spectrum of the saxophone sound was filtered using the parametric equalizer described in Sect. 4 such that the sounds deviated from the plausible monaural spectrum as provided by the filtering of the pinnae. In contrast to the audio engineering tradition, the tonal balance was secondary to the spatial position of the auditory events, since the latter is crucial for the reception of the piece.
- In some instances, the saxophone sound was band-pass filtered according to the principles of the “Blauert Bands” to move it along the Median plane or parallel to it. For example, the signal was filtered with a biquad filter at a center frequency of $8\,\mathrm{kHz}$ (12-dB enhancement) to let it appear from above (elevation effect). This way, it was possible to shift the signal from the cranial center to the front (12-dB enhancements at 250 and $3500\,\mathrm{Hz}$), rear (12-dB enhancement at $1000\,\mathrm{Hz}$), or above, while maintaining the position of the auditory event within the dimensions of the head (intra-cranial event).

The *Intra-Cranial Spatializer* was created in Pure Data to control the spatial attributes of the intra-cranial auditory events in real-time (Fig. 13). The lateral position of the auditory events can be adjusted through ILDs—either by hand (slider) or through external low-frequency oscillators with adjustable frequencies between 0.5 and $20\,\mathrm{Hz}$. The auditory events move dynamically for low frequencies, while the perceived movement stops once the oscillation period is shorter than the sluggishness of the auditory system. In this case, the incoherent signal leads to broadening of the auditory event. The psychoacoustical phenomenon can be best understood through the ILD fluctuations of the signal, because the time-variant ILDs do not necessarily decrease the coherence[7] between the left and right ear signals, since both signals remain to be highly correlated with the phase structure of the signals being untouched.

In *Cranial Transitions*, it was also anticipated to split the saxophone sound into several auditory events. This was achieved by separating the instrument’s sound into two or three streams of which two were altered spectrally through a biquad filter. The ILDs for different streams were processed independently, usually at different oscillation frequencies. This led to the segregation of the saxophone sound into two or three auditory events. The split-up of the auditory event into several streams has been extensively explained within the field of *Auditory Scene Analysis* (ASA) [6], and similar techniques have been applied in music compositions before. Roger Reynolds [28], for example, separated a single orchestral instrument into two separate auditory events by splitting the recorded signal and then modifying the relevant ASA cues in only one of two stereo channels (e.g., by applying a vibrato).

[7]The term *coherence* is seen here in context to the cross-correlation function. Blauert [5] uses this term more generally to describe any differences between both ear signals and not only those that affect the interaural cross correlation function.

6. CREATION OF CRANIAL TRANSITIONS

Today, binaural recordings are often achieved by convolving an anechoic recording with the HRTFs for the left and right channels. This method was not viable for this project, because this technique typically cannot projects sound sources in the near-field of the head, due to the circumstance that measurements of close-field HRTFs [8] have been rare so far. Sound sources close to the ear are detected fairly easily by the auditory system, because only here large ILDs are observed in the low frequency range.

The simulation of close distances was very important to obtain seamless transitions between the intra- and extra-cranial events. The following procedure was established to convert an extra-cranial event into an intra-cranial event. The recording started in the far field of the dummy head at approximately $3\,\mathrm{m}$ distance from the head’s center. While playing, the saxophone was moved toward one ear of the dummy head until it was only about $0.5\,\mathrm{m}$ away from it. The saxophone was kept in this position during the recording of the intra-cranial parts.

The subsequent methods were used to spatialize the sound intra-cranially. Firstly, the sound was converted to mono by taking only the ipsilateral channel, to eliminate the natural interaural cues and increase the correlation between both ear signals. At this stage, the equalization filter as described in Sect. 4 was applied as well. To simulate a smooth transition into the left or right side of the head, an initial ILD was applied to the mono-converted signal to make it appear at the left or right edge of the head. By steadily removing the ILD, it was possible to move the auditory event into the cranial center.

7. TIME STRUCTURE

Cranial Transition starts with atonal figures from the far field at about -30° azimuth (front-right). The recording is virtually placed in a relative small room ($1.0\,\mathrm{s}$ reverberation time). From the beginning, the sound source moves gradually toward the near field of the right ear. This effect is achieved by both physically moving the instrument toward the dummy head during the recording and gradually reducing the direct-to-reverberant energy ratio in the convolution engine by reducing the reverberation level. At 0’32”, the sounds flips into an intra-cranial event as described in Sec. 5. Initially the sound source is lateralized fully to the right, but starts soon moving toward the center, and the saxophone is manipulated using the *Intra-Cranial Spatializer* to separate the sound into multiple laterally moving streams.

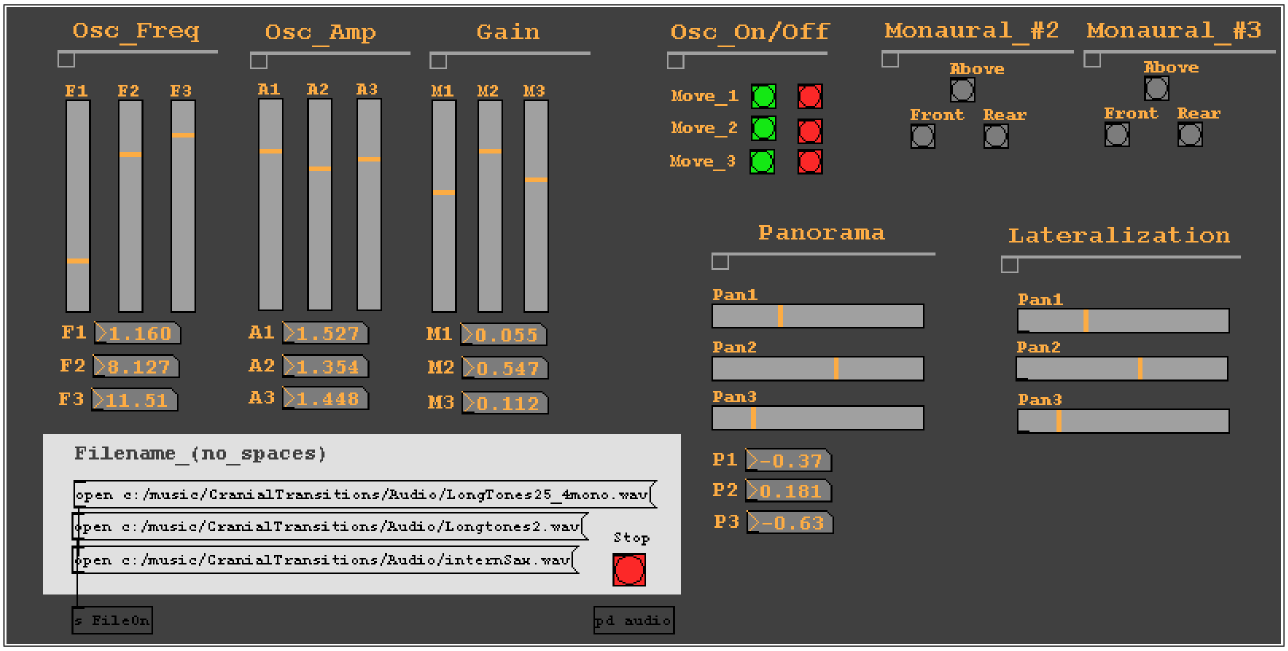

Figure 13: Graphical User Interface (GUI) for the *Intra-Cranial Spatializer*.

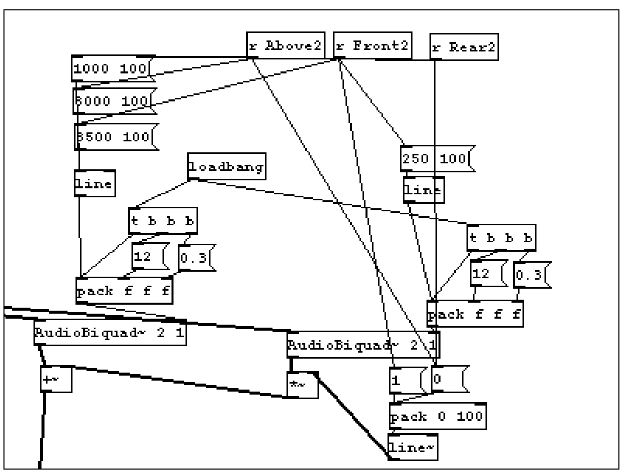

Figure 12: Pure Data patch to generate monaural cues.

At 1’28”, the saxophone is projected again as an extra-cranial event, this time from the left side. For this part, the virtual room was enlarged (2.0 s reverberation time), while keeping the wall absorption coefficients from the last setting. Again, the sound moves gradually into the near field of the ear, and transforms into an intra-cranial event at 2’38”. This time, the intra-cranial events are processed with dynamic ILDs from the beginning, and a bandpass filter with a center frequency of 8 kHz is frequently applied to project the events at the top of the head. This passage is characterized by continuous phrases with circular breathing and multiphonics, representing a restless disoriented mind that often lies below the surface.

After a brief pause, the piece continues at 3’53” with extra-cranial events in a concert hall setting (2.0 s reverberation time). While the phrasing is still continuous (circular-breathed), the phrasing is more relaxed and melodic—just as we often like to present ourselves to the outside world. In the next four minutes, the piece changes twice to an intra-cranial event space: 5’54”–6’52” and 8’03”–8’54”.

The movements of the saxophone sounds during the time interval from 8’54” to 10’27” were created through self-rotation while playing in front of the dummy head. In consistency with all other externally projected parts, no electronic processing was involved in this part of the piece, except for the impulse-response convolution.

Later at 10’27”, the piece switches continuously between the intra-cranial and extra-cranial projection spaces until the piece ends in 12’06” (see Fig. 14).

A link to the recording of *Cranial Transitions* is available at: http://www.jonasbraasch.com

8. REFERENCES

[1] Allen J B, Berkley D A (1979) Image method for efficiently simulating small-room acoustics. J Acoust Soc Amer 65, 943–950

[2] Blauert J (1969) Untersuchungen zum Richtungshören in der Medianebene bei fixiertem Kopf. Doct diss. Rheinisch-Westfälische Techn. Hochschule Aachen, Aachen

[3] Blauert J. (1969/70) Sound localization in the median plane, Acustica 22:205–213

[4] Blauert J, Lindemann W (1986) Auditory spaciousness: Some further psychoacoustic analyses, J Acoust Soc Amer 80, 533–542

[5] Blauert J (1997) Spatial hearing – the psychophysics of human sound localization. MIT press, Cambridge MA

[6] Bregman A S (1990) Auditory scene analysis. The perceptual organization of sound, Cambridge, Mass., MIT Press

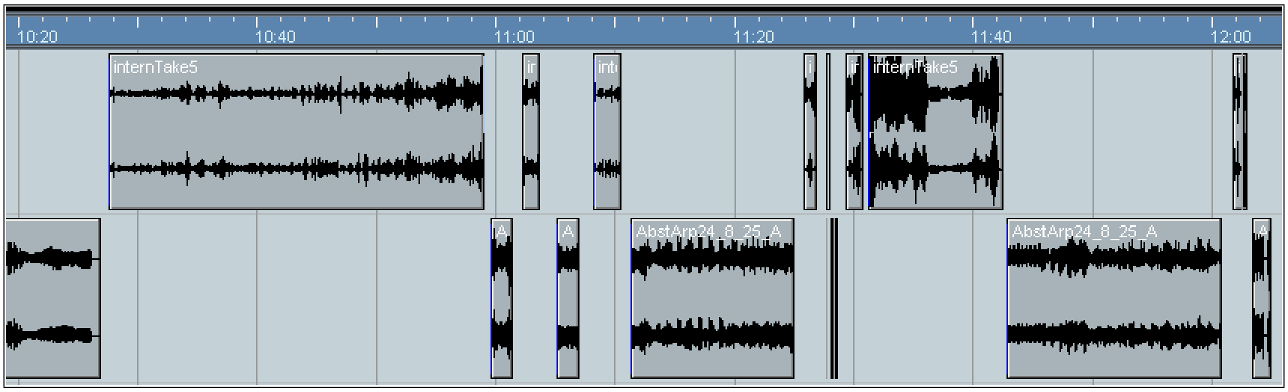

Figure 14: Timeline for the end of *Cranial Transitions*, time interval: 8'54" to 10'27". The top track projects intra-cranial events, the bottom track extra-cranial events.

[7] Braasch J, Hartung K (2002) Localization in the presence of a distracter and reverberation in the frontal horizontal plane. I. Psychoacoustical data, ACUSTICA/acta acustica 88, 942–955.

[8] Brungart D S, Rabinowitz W M (1999) Auditory localization of nearby sources. Head-related transfer functions. J Acoust Soc Amer 106, 1465–1479

[9] Chernyak R I, Dubrovsky N A (1968) Pattern of the noise images and the binaural summation of loudness for the different interaural correlation of noise. Proc 6th Int Congr Acoustics Tokyo, Bd 1, A-3–12

[10] David E E, Guttman N, von Bergeijk W A (1959) Binaural interaction of high-frequency complex stimuli. J Acoust Soc Amer, 31:774–782

[11] Deutscher Normenausschuß (ed.): Schallabsorptionsgrad-Tabelle. Beuth-Vertrieb GmbH, Berlin, 1968.

[12] Fletcher N H (1940) Auditory patterns, Rev Mod Phys 12:47–65

[13] Gaik W (1993) Combined evaluation of inter-aural time and intensity differences: Psychoacoustic results and computer modeling. J Acoust Soc Amer 94:98–110

[14] Goldberg J M, Brown P B (1969) Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: Some physiological mechanism of sound localization. J Neurophysiol 32:613–636

[15] Harris G G (1960) Binaural interaction of impulsive stimuli and pure tones. J Acoust Soc Amer 32:685–692

[16] Hartmann W M, Wittenberg A T (1996) On the externalization of sound images. J Acoust Soc Amer 99, 3678–3688

[17] von Hornbostel E M, Wertheimer M (1920) Über die Wahrnehmung der Schallrichtung, in: Sitzungsber Akad Wiss, Berlin, 388–396

[18] Joris P J, Yin T (1995) Envelope coding in the lateral superior olive. I. Sensitivity to inter-aural time differences. J Neurophysiol 73:1043–1062

[19] Joris P J (1996) Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. J Neurophysiol 76:2137–2156

[20] Knufinke C, SIR Impulse Response Processor, http://www.knufinke.de/sir/

[21] Kuhn G F (1977) Model for the inter-aural time differences in the azimuthal plane. J Acoust Soc Amer 62:157–167

[22] Licklider J C R (1948) The influence of interaural phase relations upon the masking of speech by white noise. J Acoust Soc Amer 20:150–159

[23] Litovsky R Y, Colburn H S, Yost W A, Guzman S J (1999) The precedence effect. J Acoust Soc Amer 106:1633–1654

[24] Mills A W (1958) On the minimum audible angle. J Acoust Soc Amer 30:237–246

[25] Patterson R D, Allerhand M H, Giguère C (1995) Time-domain modeling of periphal auditory processing: A modular architecture and software platform. J Acoust Soc Amer 98:1890–1894

[26] Puckette, M, Pure Data, a patchable environment for audio analysis, synthesis, and processing, with a rich set of multimedia capabilities, http://www-crca.ucsd.edu/~msp/software.html

[27] Lord Rayleigh (1907) On our perception of sound direction, Philos Mag 13:214–232

[28] Reynolds R (2002) Title, form and method. Composing music. The Rothschild essays, (S.McAdams, ed.), New York, London, Routledge

[29] Sayers B McA (1964) Acoustic-image lateralization judgement with binaural tones. J Acoust Soc Amer 36, 923–926

[30] Steinhauser A (1877) The theory of binaural audition, Phil Mag 7:181–197 & 261–274

[31] Thompson S P (1877) On binaural audition, Phil Mag 4:274–276

[32] Thompson S P (1882) On the function of the two ears in the perception of space, Phil Mag 13:406–416

[33] Toole F E, Sayers B McA (1965) Inferences of neural activity associated with binaural acoustic images. J Acoust Soc Amer 37, 769–779

[34] Woodworth R S, Schlosberg H (1962) Experimental Psychology, 349–361. Holt Rinehard Winston, New York

[35] Zahorik P (2002) Assessing auditory distance perception using virtual acoustics. J Acoust Soc Amer 111, 1832–1846