

## STAX SR-007 (OMEGA II) ... A REVIEW AFTER 4 YEARS OF OWNERSHIP

*A madman is one who 'hears voices in his head'.*

*Headphone-user: 'you calling me a madman?'*

In September 1999 I posted a review of STAX SR-007 (Omega II) headphone at HeadWize.

I wrote as detailed and exhaustive a review as I could manage at the time, with the target audience being obsessive headphone-users who had, like me, noticed the strange addictive joy of being immersed in a small cranium-bound soundfield that oddly pulsates with life; a soundfield that for all its smallness paradoxically triggers our imagination to 'see' large acoustic spaces.

Four Christmases have passed since that 1999 review. Pleasant and unpleasant things have happened in my personal life during these past years; the chief unpleasant thing being the sheer fact that I have aged 4 years, and the chief pleasant thing (so I console myself) being that I have grown four years wiser.

(Just in case you are wondering, I am 38 now.)

At the time of my 1999 review, I thought that the number of interested readers could be counted with one hand - back then I just didn't think that there were that many people interested in high-end headphones who also frequented headphone forums. Today I am gleeful to see how many fellow headphone enthusiasts there are out there, judging from the activity here at Head-Fi. I am also quite amazed to observe how many owners of high-end headphones and high-end amps there are who are presently "visible" in the forums, compared to the scant few back in 1999.

I have disappeared for a long time from HeadWize / Head-Fi. I found that writing a post, especially a full-length review, to be quite consuming, which was another reason why I stopped posting for a long time. It's simply far more relaxing to disappear from the forums and enjoy my headphones. But lately I came back as a forum lurker, and have enjoyed reading dozens and dozens of threads. There are many intelligent members here, and I was entertained and educated by the experiments, insights and exchanges (some heated) posted by headphone enthusiasts from all over the world. The folks who go back to HeadWize days may remember me - I suspect most people here at Head-Fi either don't know me or know me only as a ghost from the past. I wish to say hi to everyone.

### **WHICH READER AM I ADDRESSING?**

Because a headphone forum comprises of different people with all sorts of headphone experience levels and all sorts of listening habits, I had to be clear in my mind for whom I was targeting this write-up.

This essay is rather detailed, and unfortunately may be difficult to digest. I have tried my best

to sequence the flow of this essay such that the reader is gently, gently eased into increasingly complex concepts. But what I've not done is to dumb down the essay. I resisted the urge to simplify the concepts because I do not want to short-change those readers who are highly curious about I have to share here.

Readers who listen predominantly to close-miked music (such as rock and pop) may find the concepts rather alien and detached. Headphone-users who listen predominantly to close-miked music are more apt to go "so what?" or worse "what bullshit is this?" to a large part of this article, because the things mentioned here lie outside of their scope of experience. If this describes you, I hope you can suspend disbelief just for the duration of this article, so that the knowledge gained from this write-up would lie dormant in your memory. In some future moment when you least expect it, you hear something either at home or at the audio shop (or at a Head-Fi Meet perhaps?) that will remind you of what you read here.

Readers who habitually listen to music with a lot of ambient cues (such as live jazz, orchestral and choral) will more readily understand how the spatial subtleties mentioned in this write-up relate to headphone listening. Such readers may have less problems diving into the intricacies elaborated later on.

Readers of my review of the Omega II written 4 years ago may remember that I have used the term "headstage" before, but I did not manage to explain its meaning clearly in that review - hence some readers may have been puzzled as to the purpose of its inclusion then. I apologize for your warranted puzzlement. In this current write-up I have finally succeeded in nailing down the meaning of "headstage" in no uncertain terms. Additionally, I have found a way to explain the Four Depth Cues in a clear and communicative manner. (The Four Depth Cues first appeared in my archived essay at HeadWize's Library, but this current write-up takes it one step further by having a headphone review structured on the Four Depth Cues.)

It has taken me years to crystallize these concepts into a consistent framework. I am happy to share with you today the fruits of my labour.

### **OBJECTIVES OF THIS ARTICLE**

The objectives of this write-up are twofold:

Objective 1: to share my feelings of the STAX SR-007 (Omega II) after 4 years of ownership. Am I still happy with my purchase, now that the new-toy-syndrome has passed? A comprehensive review of a product owned after a passage of time must surely furnish a better indication to another prospective buyer of that product's worth (or lack thereof) than a review written during the honeymoon period. Also, it is fiendishly difficult to accurately describe the sonic character of a headphone - any headphone. A few of my detailed observations now differ from those I made in 1999. Back when I was active in the forum, there were instances where I promoted this headphone as the best headphone in the world. But today, as a jaded forum lurker, I wonder about the fruitfulness and sensitivity of such

claims. There are so many marvellous headphones out there - with a fan base for each of them - why tell others that one and only one headphone is the best? Is there such a thing as a single best headphone for everyone anyway?

Objective 2: to persist in an even bigger project of mine, which is to attempt to advance the development of an adequate language to describe the sound of headphones. The language we use today has evolved through the decades within the context of a loudspeaker-centric audio world. A language specifically for headphones has not yet been constructed. Some Head-Fiers construct DIY amps - I construct here a DIY language. This is an ambitious project; one that I started 4 years ago, and it is heart-warming to see that a few people have begun to use the term "headstage" since its introduction back in 1999. In this write-up, I will be offering a crystal clear explanation of the term "headstage", and then I will be adding even more words to the lexicon of headphonespeak.

This write-up is therefore not just a simple review of the Omega II - it is also about the creation of a new language, new terminologies and a new review methodology. My review of the Omega II may at first appear sporadic and strewn all over this essay, but actually there's a structure: every time a new term has been properly defined and explained, I will subsequently proceed to review the Omega II using the newly created terminology. Then I will proceed to the second terminology, define what the new word or words mean, and proceed to describe the Omega II using the second set of new words ...and so on.

Let's start.

### **BEFORE THE FOUR COMES THE ONE**

First there is the One; then there are the Four.

I will be touching on the Four Depth Cues towards the middle of this essay, but from the beginning I want to say that there is one sonic mechanism that overrides the Four Depth Cues. This One is the sense of *sound localization*.

We acquire the sense of sound localization because our left and right ear each receives a slightly different input, and by comparing the two our brain interprets the location of the sound source. When we put on our headphones, the headphone transducers are positioned very near our ears - we can locate the source of the sound, and we are aware of this proximity of the sound source. Every time I use the word 'locate', I am referring to this One mechanism - the mechanism of sound localization. This One mechanism is more powerful than the Four Depth Cues.

This One mechanism gives rise to the *headstage*.

## **INTRODUCING: THE HEADSTAGE**

I am listening to a section of Beethoven's Pastoral symphony (andante movement), and I think there are 20 musicians packed inside my head. Listening to music via headphones can be a paradoxical experience. I know that 20 people cannot fit into my head, empty as I sometimes swear it may be during my stupider moments. Yet the steadfast illusion right now is that there are 20 musicians in my head.

There are some recordings that make me go "wow, what a huge soundstage". But here's the rub: I happen to have a wall-sized mirror on one side of my listening chair. When I look into the mirror, the illusion of the huge soundstage is stripped away and revealed for what it truly is: a cramp head-hugging soundfield. In the mirror I can "see" all those sonic images sticking to my scalp like a bad hair-do. I look away from the mirror, close my eyes, lose all sense of scaled reference to the real world, re-invest my concentration into the music, and the huge soundstage re-appears. But when I open my eyes and look again at the reflection of my headphones in the mirror, I once again "see" the scalp-bound soundfield.

I call this soundfield that stubbornly refuses to take leave of my head the *headstage*.

The difference between soundstage and head-stage is illusion and reality. The soundstage is the (desired) illusion; the headstage the (unfortunate) reality.

Another way of stating the difference between headstage and soundstage: headstage is about the localization of sonic images *in relation to your head*. Let's say you are listening to a piece of music that contains 3 sonic images. One image is located at the right temple of your forehead, another image is skimming the top centre of your scalp, and yet another image is located an inch beyond the left earcup. The arena within which all these sonic images are located is called the headstage. And it is a tiny arena - I estimate this arena on the Omega II to be maybe 8" wide and 5" tall (it could be bigger on your headphone - I've always said that the Omega II has a small headstage - but more on this later). The sound-stage is something else altogether. The sound-stage is the qualitative perception of ambient cues captured in the recorded music. The soundstage can be very big, as big as a cathedral nave, if that was what was indeed captured in the recording.

When listening to headphones we can choose between perceiving the soundstage or perceiving the headstage. Your mental concentration can swing the perception one way or the other. During moments when we are utterly absorbed in the recording, all you have to do is to tell yourself to "snap out of it", and chances are that you will "lose sight" of the majestic soundstage. What's so majestic when you choose to become aware that the whole violin section of a grand and majestic orchestra is only 4 inches wide across your forehead?

When listening via headphones, most of us choose to be aware of the soundstage instead of the headstage, in an effort to distract ourselves from noticing the cramp head-hugging soundfield or in an effort to lose oneself in the recording - the latter is valid and is after all the whole point of listening to music. But distracting yourself from scrutinizing the head-hugging

soundfield will not make you a more discerning listener. You have to understand the head-hugging headstage first, cramp as it may be, before you understand the soundstage.

### **HEADSTAGE: ANALOGY OF A PHOTOGRAPH**

What is the headstage, really? First I will put forward an analogy, then I will offer a working definition of the term “headstage”.

Analogy: imagine a 5-inch wide photograph depicting a sprawling mountain scene going on for miles and miles. A photograph is nothing more than colour pigments distributed on a flat piece of paper. There is no mountain on the piece of paper, nor inside nor behind the piece of paper. The mountain is in the eye of the beholder. Furthermore, a photograph does not need to be mountain-sized in order to depict a mountain. Additionally, a statement that the mountain in the photograph is 10 miles away does not contradict the fact that the colour pigments representing the mountain are lying flat on a piece of paper.

The two-dimensional headstage is analogous to the two-dimensional photograph. If a small photo can depict a large scenery, why can't a small headstage portray a large soundstage? And if a flat photo can depict distance, why can't the two-dimensional headstage depict depth?

This is the definition of the term “headstage”:

**the headstage is a flat plane, small in size, positioned vertically such that the plane intersects both ears, and all sonic images are chained to the two-dimensionality of this plane.**

None of my past articles has offered such a concise definition of “headstage”.

Please take time to digest this: *all sonic images are chained to the two-dimensionality of the headstage, much the same way the mountain is chained to the two-dimensionality of the photograph.*

Why do I say that the headstage is two-dimensional? In order to be aware that this head-hugging soundfield is actually two-dimensional, you have to stop yourself from being swept away by the soundstage illusion of the recording, and start to focus on the location of the images in relation to your head. Your headscape offers several landmarks that you can reference the location of the images against. Landmarks on your head include the front centre of your forehead between the eyebrows, the front centre of your forehead where your third eye would be if you were a Buddha, front top of your forehead where your hairline is if you haven't started balding yet, the left and right temples of your forehead, and the left and right ears on your head. It may seem unnatural at first, but try not to focus on the soundstage cues inherent in the recording, but instead focus on the location of images in relation to your headscape.

Then you will realize the truth that all the images can be located more or less on a flat vertical plane. Average playback systems will create flatter sonic images that resemble stickers from a child's sticker book. Sonic images are like flat stickers that you can "paste" on the flat vertical headstage. Superior playback systems create more rounded, full-bodied images, in which case the headstage resembles more an upright rectangular tupperware\* within which all sonic images are contained. (\*tupperware = plastic food container, just in case there's a cultural gap here.) But whether it is a flat plane or an upright tupperware, the point here is that whilst there is depth in the recording, there is no depth to the *localization* of the images.)

I have read accounts of a headphone's soundfield as being "a clothesline stretched from one ear to the other", or another account describing it as being "three blobs in the head". My senses tell me that both descriptions of the headstage shape are inaccurate.

I simply don't perceive the images being located as if they were strung along a straight line going from ear to ear, like so many beads on a string. There is such a thing as height, so the one-dimensional description of the headstage is something that contradicts my personal experience. A straight line going from ear to ear is actually located very deep in my skull (a straight line going from ear-to-ear is three inches below the top of my scalp) and the only time I noticed images located three inches below the top of my scalp is when I listened to mono recordings. Stereo recordings create not just left-to-right differentiation, but also create a sudden upward expansion of the headstage, i.e., the creation of headstage height. (If you have a Stereo-Mono toggle switch on your amp you will notice that toggling to Mono will collapse the headstage into a tight-fisted ball deep inside your head, while toggling to Stereo will not only provide left-to-right differentiation but also expand the headstage *upwards*) So the description of a headstage as a thin clothesline stretching from ear to ear is something I take issue with.

As for the description of the headstage as being "three blobs in the head" - on my systems (past and present) I have not heard the three blobs effect. Intellectually I understand what HeadRoom is trying to say - it's just that the three blobs effect simply doesn't square with what I have experienced so far. I suspect that HeadRoom offered such a stark model (three blobs is a very stark model) because a more subtle explanation of the crossfeed mechanism may potentially be lost on laymen. In an advertisement, you need a clear, strong message; and the three-blobbed headstage is as clear a message as you can get: "*you don't want the three blobs - you want our crossfeed*". From my experience, the headstage is a smooth continuum from left to right; and there is no distinct separation into three separate blobs, unless I was playing a very old stereo recording - as old or older than myself. (This is not to be construed as a comment on the crossfeed mechanism. I am commenting on the accuracy of the description of the headstage as being a three-blobbed affair.)

I am prepared to accept a description of the headstage shape as being a spherical soundfield, but it is a *squashed* sphere, more like an oblong rugby ball: the left-to-right dimension is larger than the front-to-back dimension. A person who insists that the

headstage soundfield is a perfect sphere must either get his ears checked or tell us all what super-duper headphones he is using that can create not only left-to-right localization but front-to-back localization as well. (Binaural recordings that matches one's personal HRTFs and various 3D-processing methods lie outside the scope of this write-up. This write-up is restricted to stereo headphones playing stereo recordings.)

The description that most resembles my experience of the headstage shape is any one of the following: that it is either a flat vertical plane or an upright rectangular tupperware or an oblong-shaped ball or a thick fat discus placed vertically. Whatever shape you choose to describe the headstage as, the main thing is that this shape has a *larger left-to-right dimension and a very flat front-to-back dimension*. (But if I were to be absolutely accurate about it, I'd say that the headstage is a rainbow-shaped arch springing from ear to ear with the apex of the rainbow at the top centre of the forehead. All images are located in a smooth continuum along this rainbow. This rainbow has a larger left-to-right dimension and a very flat front-to-back dimension.)

Most headphones create headstages that intersect the ears. (Meaning to say that the vertical plane or the oblong ball or the upright tupperware or the vertical discus or the rainbow intersects the ears.)

But headphones such as AKG K1000, STAX SR-Sigma and -Sigma Pro create headstages that do not intersect the ears but instead their headstages are located perceptibly more towards the front. I am not so familiar with the K1000, but for the Sigmas the headstage is about 2 inches in front of the forehead. This is because their transducers are, by design, angled perpendicularly and located more frontally than in other headphones.

This is where I review the Omega II for the first time in this essay. What about the Omega II's headstage?

The Omega II's headstage does not intersect the ears, but is located very slightly in front, such that the headstage is in contact with the flat front of my forehead. I guess this slightly frontal position of the Omega II's headstage (not as frontal as in the Sigmas though) is due to the headphone's slightly tilted diaphragms, such that the headphone co-opts the ear flaps at an angle, instead of directly firing the sound straight into the ear canal.

The second thing about the Omega II's headstage is that the sonic images are so rounded and full-bodied, such that the headstage does not seem like a flat vertical plane, but more like an upright rectangular tupperware into which all sonic images are contained. The longer side of the rectangular tupperware is touching the flat front of my forehead. (The tupperware is not hovering outside my forehead - the tupperware overlaps and protrudes into the front portion of my head. The frontal lobe of my brain is contained in this hypothetical tupperware.)

The third thing about the Omega II's headstage is that it is small; shockingly smaller than all headphones I remember hearing. Believers of a 'bigger is better' worldview may be in a rude shock.

The fourth thing about the Omega II's headstage is the precise way it locates sonic images within the headstage. Its headstage is small, but it can paradoxically hold a great many sonic images without seeming overcrowded. The images are located very precisely in the headstage - sometimes you feel as if the images are merely millimetres apart from each other within the headstage, but because of the awesome resolution power of this headphone, mere millimetres is enough to separate those two images.

We have come to the end of the section on "headstage". I hope you feel that the explanation offered about what the headstage is has been insightful. The way headphones erect their headstages has so far been conspicuously absent from the literature of headphone reviews. I feel that a review of a headphone - any headphone - becomes more thorough and complete when the reviewer comes to grips with these 4 things: headstage size, headstage fullness, headstage frontality (or lack of) and precision of image location within the headstage. All 4 things are about the One mechanism of sound localization.

But would the term 'headstage' be useful in every headphone review? Perhaps not. The description of the Omega II's headstage is important because its headstage is highly peculiar - small but highly focused, slightly frontal and full-bodied - these four characteristics are peculiar. Many headphones do not exhibit all four characteristics simultaneously. If headphone X's headstage is unremarkable (meaning its headstage is normal-sized and is not frontal) then it may not be necessary to describe headphone X's headstage in a review, other than perhaps a passing remark that its headstage is that normally expected of a headphone.

One further question about the headstage remains. If all sonic images are chained to the two-dimensionality of the headstage, then what gives rise to the illusion of depth? Or to rephrase the question: how does one reconstruct soundstage depth from the two-dimensional headstage?

### **INTRODUCING: FOUR DEPTH CUES**

The Four Depth Cues are the mechanisms by which the two-dimensional headstage is given a semblance of the third dimension. These Four Depth Cues transform the headstage into the perceived soundstage. The photograph analogy is once again helpful here.

Let's assume that you are looking at a photograph that depicts both nearby mountains and faraway mountains. How do you know that certain mountains in the photograph are closer to you whilst other mountains in the same photograph are further from you? The photograph is a flat piece of paper - but it communicates depth cues via five visual cues:

Visual cue 1 - mountains or objects that are small in the photo may be interpreted as being far, unless otherwise contradicted by other cues

Visual cue 2 - mountains with lighter colour in the photo may be interpreted as being far, unless otherwise contradicted by other cues

Visual cue 3 - mountains in the photo that have more terrain detail appear nearer, unless otherwise contradicted by other cues

Visual cue 4 - mountains seen through an atmospheric haze in the photo appear far, unless contradicted by other cues

Visual cue 5 - a mountain that overlaps and blocks another mountain in the photo is perceived as being the nearer one, and this visual cue takes precedence over all other visual cues

The above are the five mechanisms that afford visual depth cues in a photograph. The mechanism of perceiving distance operates thus: TWO-DIMENSIONAL PHOTO - -> FIVE MECHANISMS OF VISUAL CUES - -> PERCEPTION OF DISTANCE (DESPITE THE FLATNESS OF THE PHOTO).

For each of the above visual cue there is a corresponding sonic equivalent. I will re-list the five visual cues, but for each visual cue I will now provide its sonic equivalent:

Visual cue 1 - mountains or objects that are small in the photo may be interpreted as being far, unless otherwise contradicted by other cues

**Depth Cue #1- sonic images that are softer in volume appear further, unless otherwise contradicted by depth cues #2, #3 and #4**

Visual cue 2 - mountains with lighter colour in the photo may be interpreted as being far, unless otherwise contradicted by other cues

**Depth Cue #2- sonic images that sound tonally attenuated appear further, unless contradicted by depth cues #3 and #4**

Visual cue 3 - mountains in the photo that have more terrain detail appear nearer, unless otherwise contradicted by other cues

**Depth Cue #3- sonic images that have more textural detail appear nearer, unless otherwise contradicted by depth cue #4**

Visual cue 4 - mountains seen through an atmospheric haze in the photo appear far, unless contradicted by other cues

**Depth Cue #4- sonic images swathed in a diffused/reverberative halo appear further**

Visual cue 5 - a mountain that overlaps and blocks another mountain in the photo is perceived as being the nearer one, and this visual cue takes precedence over all other visual cues

**There is no sonic equivalent to this mechanism because sonic images are "transparent enough" such that one sonic image cannot "block" another**

The above are the four mechanisms that afford sonic depth cues in a headstage. I call these the Four Depth Cues. The mechanism of perceiving distance operates thus: TWO-DIMENSIONAL HEADSTAGE ---> FOUR DEPTH CUES ---> PERCEPTION OF

## SOUNDSTAGE DEPTH (DESPITE THE FLATNESS OF THE HEADSTAGE).

Please note that these Four Depth Cues do not free the images from the bondage of the head-stage. The images are still chained to the head-stage plane, just like the way the faraway mountains and nearby mountains are still chained to the two-dimensionality of the photograph. The mechanisms only offer the facsimile of depth, but not real depth itself. *The Four Depth Cues do not create out-of-the-head images.*

For purposes of layout clarity I will re-list the Four Depth Cues here:

**Depth Cue #1- sonic images that are softer in volume appear further, unless otherwise contradicted by depth cues #2, #3 and #4**

**Depth Cue #2- sonic images that sound tonally attenuated appear further, unless contradicted by depth cues #3 and #4**

**Depth Cue #3- sonic images that have more textural detail appear nearer, unless otherwise contradicted by depth cue #4**

**Depth Cue #4- sonic images swathed in a diffused/reverberative halo appear further, and this cue takes precedence over all other cues**

You will notice that there is a ranking order to the four cues, starting with #1 as the weakest of the four cues and #4 as the strongest of the lot. This hierarchical order was arrived at after careful observations by listening to many recordings via my headphones over the past 8 years.

I will now explore each of these four cues in detail. For each of the four cues I will also touch on qualities of the audio playback chain (source-amp-headphone) necessary for the accurate portrayal of that respective mechanism. I will also review the Omega II's ability to render each of the mechanisms.

### **DEPTH CUE #1:**

**sonic images that are softer in volume appear further, unless otherwise contradicted by depth cues #2, #3 and #4**

Hypothetical scenario: You are in the middle of a losing cavalry battle. Hope is almost lost, but out of the blue you hear a bugle call from afar: friendly reinforcement is approaching. Suddenly there is hope that you can save your cavalry division from certain defeat. Something so soft-sounding as the bugle call from afar has stirred intense feelings of hope.

Great depths of romantic feelings can be ascribed to the soft-sounding sonic image, and there are many instances in recorded music of all types where you find the soft-sounding sonic image being the prime carrier of emotion and meaning during that particular musical passage.

(Psychoacoustically, we interpret the soft-sounding image to be far away because we have learnt from infancy that an object making a sound or noise will sound softer as the object moves further from us.)

The challenge that the soft-sounding sonic image poses to the audio playback chain is this: how do you sustain the presence of the soft-sounding image amidst all the other louder sounds? How do you prevent it from being drowned by those louder sounds? Even more difficult: as those loud sounds alternate between being loud, being soft and being even louder, how do you prevent the soft-sounding image from flickering in and out of existence at the mercy of those fluctuating loud sounds?

The challenge posed here to the audio playback system is therefore one of clarity and resolution, and to a lesser extent, one of macrodynamics. A system with sufficient clarity will differentiate the soft-sounding image from the louder images. Systems with good portrayal of macrodynamics would allow the various instruments to go loud or soft, and in superior playback systems, the instruments will go louder or softer independently of each other.

The other challenge to the audio playback system is how to tell if the image is soft because it is far away, or because it is deliberately played softly by a nearby musician. The latter retains textural intensity but not volumetric intensity. (Textural intensity is touched on in the section on Depth Cue #3.)

How well does the Omega II fare in the rendition of the First Depth Cue (#1)?

In a word: stupendous. This headphone is capable of oodles of detail, and the soft-sounding image never gets lost even in a cacophonous jungle of other loud sounds. Recordings that rely on #1 for the portrayal of depth cues are rendered vividly. Image stability of the soft-sounding image is extremely high.

As an example, I am now listening to the soundtrack from *Mighty Joe Young*. The beginning of Track 2 has a soft-sounding image of a piano tuned weirdly (ala John Cage-like), played percussively but very softly, and its softness gives the impression that it is further away compared to the louder percussive slapping of sticks and the soaring of violins. On the Omega II, the image stability of this soft-sounding image is maintained despite the fluctuations in volume of the louder sonic images.

Another example: Princess Leia's Theme from the soundtrack of *Star Wars*. This is a sweet, lovely slow piece, with a solo flute opening the track, followed by a solo clarinet, then a solo horn takes up the main theme. When the solo horn is carrying the main melody, a background violin provides the accompaniment. The violin is played softly as well as played a little further away. The softness (#1) and lack of textural specificity (depth cue #3) of the violin provides the depth and backdrop to the perceived acoustic space, whilst the louder and more texturally specific solo horn is the foreground object. The solo horn presents a high image height - as a foreground object it "stands tall" in the acoustic space. (That's the lovely thing about horns and human voices - whether solo or massed - they tend to "stand tall" in

the acoustic space.) Princess Leia's Theme develops slowly but inevitably to its mournful conclusion - at the end, a solo violin weeps its last farewell note, gently dying into the night. (With such a sweet but sad ending to the theme, it's a wonder that the Princess didn't die in the movies.) The Omega II convincingly portrays the layered perspectives of this theme utilizing depth cue #1 (as well as #3 - but more on this later).

But if a sonic image is soft-sounding, couldn't it be that the instrument was played softly by the musician and not because the instrument was far away? How do you differentiate between the two? This is how: on a hierarchical order, depth cue #1 is at the bottom of the rung, and can be overridden by depth cues #2, #3 and #4. Depth cue #1 is the weakest of the four cues. You will perceive a volumetrically soft image as being far away, per depth cue #1. But if you hear a volumetrically soft but tonally rich image, #2 will override #1, and you perceive the volumetrically soft image to be nearer.

Example: I am now listening to Stravinsky's The Soldier's Tale (Track 6 The Three Dances). The track opens with a violin and timpani, then a soft-sounding gentle cymbal crash from the rear of the stage. Or at least the soft-sounding cymbal seemed at first listen to come from the rear of a deep stage, due to the effects of depth cue #1. But on closer listen, the cymbal was in fact played softly rather than played faraway. How can I tell? Because while a faraway cymbal would lose much of its metallic shimmer via depth cue #3, the soft cymbal crash I heard in this track retained a highly specific metallic shimmer. (In talking about the texture of an instrument I have actually gone a little ahead of myself. Textural specificity as a depth cue is touched on later when I come to Depth Cue #3.) This soft-sounding cymbal crash retained too much texture for it to be far away - implying that it is nearby. High-end headphones like the Omega II make it easier to differentiate between those two situations.

Another example where the Omega II allows me to experience depth cue #3 overriding depth cue #1: Death Of Darth Vader (a fellow Sith, by the way), from the soundtrack of Return Of The Jedi. Towards the ending of this piece, when Vader dies in his son's arms, a gently plucked harp softly plays Darth Vader's Theme. (Usually Darth Vader's Theme is pompous and militaristic, played by snare drums and brass instruments; but in this scene where he dies, a harp - a harp! - takes up the theme.) The softly plucked harp sounds unmistakably near despite depth cue #1. The leading edge textural detail of the plucked harp is clearly heard - I can almost "see" the fingers plucking the harp strings. Depth cue #3 says that when the textural detail is high, we perceive the image to be near. We can infer from this observation that depth cue #1 is easily overridden by depth cue #3.

### **DEPTH CUE #2:**

**sonic images that sound tonally attenuated appear further, unless otherwise contradicted by depth cues #3 and #4**

Hypothetical scenario: You print out on hard copy the threads at Head-Fi titled "Do You Believe In God?", "In God We Trust?" and "Jude vs God". You bring the printed stack

outdoors to read, where you hope that the bright outdoor light would conspire with your reading concentration to finally put the question of the existence of God to rest.

You come across the part which goes “of course God does not exist“ when the distant roll of thunder rumbles across the sky. And then you get the hint: He exists, and has just sent you a gentle reminder. You think to yourself: He could have given you a more severe rebuke by sending forth a deafening thunder clap 10 feet from where you sit, replete with a high-pitched transient snap, like two Godzilla-sized kendo sticks forcefully meeting each other in mid-air.

But no. Instead you heard....the distant thunder rumble.

What the distant thunder roll lacks in high-pitched proximity, it makes up for in majesty, for it rumbles across the land with a deep and authoritative resonance. But how did you know the thunder was distant? (The distant thunder was still quite loud; so it was not through depth cue #1.)

You inferred that the thunder was distant because it lacked high frequency components.

Every sound, except for pure test tones, contains high frequency harmonics and low frequency harmonics. When the source of the sound is nearby, the full palette of all these harmonics can be heard together with the principal harmonic.

But in a free field, such as in the open outdoors, the further sound has to travel, the more it loses its high frequency content. Which is why thunder from afar is made up of mostly low frequency sounds. The high frequency components have been attenuated along the way.

In a diffuse field, however, such as in a concert hall, it is my observation from recorded music that as sound travels further, it loses both its high frequency and low frequency content. It is a tough call whether the low-frequency harmonics also gets attenuated, because I think it differs from one recording venue to another. It depends on the acoustic character of each venue whether or not the low frequency component also gets attenuated. It also depends on the microphone array, recording equipment and the recording artist’s decisions. (But that the high frequency harmonics gets attenuated in a diffuse field occurs so frequently in recorded music that I can take it as a given.) In some recordings, hall ambience actually comprises of low frequency harmonics.

*Sidetrack: in saying that depth cue #2 is a result of tonal attenuations, I might be putting the cart before the horse. It might actually be the opposite: we judge the tonal balance of a recording or a headphone based on how far or how near everything sounds. After all, our ears don’t behave like frequency spectrographs - we don’t plot frequency spectrums with our ears. We perceive what is far and what is near - we use expressions such as “forward-sounding midrange” and “laid back treble”. When a certain portion of the frequency spectrum consistently sounds nearer irrespective of recording, we say that the headphone has an accentuated bump in that portion of the spectrum. It is the perception of forwardness via depth cue #2 that allows us to estimate a headphone’s tonal hot spots.*

There are two incarnations how depth cue #2 manifests itself, and this depends on the recording.

First incarnation is called tonal blandness (#2a): there is a simultaneous attenuation of high frequency harmonics and low frequency harmonics. This results in the distant sonic image sounding more tonally bland. It is very satisfying to hear the effects of distance on the tonal character of instruments. It seems odd to say that it is satisfying to hear the loss of tonal richness of an instrument - shouldn't it be the opposite: that it is satisfying to hear the tonal richness of an instrument? Well, both are satisfying in their own ways. Small works that are miked more closely will give me the tonal richness and intimacy of each instrument, whereas distance-miked works will give me the satisfaction of hearing greater distances and a grander scale to the proceedings. Sometimes a sonic image is meant to be tonally bland due to effects of distance.

The second incarnation of depth cue #2 occurs when there is an attenuation of only the higher harmonics, with a preservation of low harmonics. This leads to what I call a harmonic shift (#2b), to coin a new term. When the higher harmonics are attenuated due to the effects of distance yet the lower harmonics remain largely intact, the resultant tonal character of the sonic image shifts towards the lower harmonics. The sonic image seems deeper-sounding, with more heft in the lower regions. (It's always a harmonic shift downwards - never upwards. The example of the distant thunder roll described at the start of this section is an example of harmonic shift.)

What challenge does depth cue #2 pose to the audio playback system?

The challenge that the Second Depth Cue poses to the audio playback system is two-fold: tonal neutrality and harmonic diversity, to coin a new term.

The first challenge is tonal neutrality. If the headphone is not neutral, i.e. if there are segments in the frequency spectrum that are spotlighted at the expense of others, this would create havoc to the sense of perspective afforded by the Second Depth Cue. I suspect that the headphone that portrays depth cue #2 just right is the Grado HP-1; but I'm saying this from memory. (See sidetrack below.) The second challenge that #2 poses to the audio system is harmonic diversity. Nearer images sound tonally richer, while further images sound tonally blander. You need an audio system that can portray tonally rich images and tonally bland images simultaneously. The ability to portray differing tonal richness fosters a sense of differing depths between images.

*Sidetrack: To be sure, tonal neutrality is a complex issue for headphones because almost all headphones are voiced for what is called "diffuse field equalization". Due to complexities in the coupling between earcup and ears, specific tonal adjustments have to be introduced for a headphone to sound tonally neutral. A headphone with a ruler-flat frequency response would sound awful. But I can swear there does not seem to be a single consistent execution of diffuse field equalization, because I observe that almost all headphones purporting to be diffuse field equalized sound so tonally different from each other.*

How does the Omega II fare in the rendition of the Second Depth Cue? Awesome, but with one point of weakness.

First the awesome point: the Omega II has a prodigious low frequency weight. You would never expect an electrostatic headphone to have so much heft in the bass regions. A weighty low frequency is critical to the portrayal of depth cue #2b (harmonic shift), which is one of the two incarnations of depth cue #2. The Omega II portrays harmonic shifts convincingly. For example, right now I am listening to Track 10 (The World Spins) from Julee Cruise's Floating Into The Night (which features the Main Theme from The Twin Peaks). Did you know that a cymbal, which most of us would expect to be a high-frequency instrument, can actually sometimes portray low-frequency harmonics? The squeezed-air cymbal on Track 10 sounds as if it comprised more low-frequency harmonics than high-frequency harmonics - a surprise to me when I became aware of it. That the squeezed-air cymbal sounded this deep contributed greatly to its sense of distance, via depth cue #2b (harmonic shift).

The Omega II's weakness in portraying #2?

The Omega II's overall tonal balance errs on the side of warmth. (Warm = clockwise tilt of the frequency balance about the fulcrum at 1kHz - definition from Stereophile) In other words, this headphone's treble is restrained (but much more on this in a later section of this write-up). The result of this treble-shy tonal balance is that the attenuation effects via depth cue #2 occur at a faster rate than what I suspect is accurate. We know that high frequency harmonics of an instrument gets reduced over distance (#2), but it seems to get attenuated a tad quicker via the Omega II.

### **DEPTH CUE #3:**

**sonic images that have more textural detail appear nearer, unless otherwise contradicted by depth cue #4**

Hypothetical scenario: You have been a RS-1 user for years. You swear by its clarity and textural immediacy. Your friend who owns a HD600 invites you over to his house to try out his headphones. You have never auditioned the HD600, so you trudge over to his house with a clutch full of your favourite CDs.

You go "what!?", when you finally get a handle on the HD600's character. You complain of its distant mid-hall perspective. You even complain that the HD600 sounds "veiled".

When you get back to your home, you start a new thread at Head-Fi titled "Shocking news! HD600 is veiled and distant-sounding", thereby starting yet another argumentative thread.

For the previous two depth cues, I started off with wacky scenarios to give a humorous touch to the proceedings. For the Third Depth Cue, I can think of no other better anecdote than one that involves the RS-1 and HD600, which had been the topic of many previous feuds at both HeadWize and the early days of Head-Fi. I wish to provide here a fresh angle on the differences between these two headphones.

Depth Cue #3 says that sonic images that have greater textural detail appear nearer.

The RS-1 is the more detailed headphone - it portrays more sonic information on the textures of instruments. Via depth cue #3, this creates the impression that the instruments are nearer to the listener. Depth cue #3 is the reason why we customarily say that HD600 is more mid-hall, while RS-1 is closer to the stage. One criticism of the RS-1 that I am hesitant to agree wholeheartedly with is that it is coloured - it has become too commonplace for audiophiles to accuse a component of being coloured when the only sin that that component ever committed was to be texturally specific.

(I made the same mistake 4 years ago in my review of Omega I vs Omega II, when I referred to the Audio Note DAC2 digital-analogue converter as being coloured, when what I actually meant was that this lively DAC was texturally specific. My apologies to Peter Qvotrup, who did give me a gentle rebuke on this matter and insisted that his DACs were not coloured when I e-mailed him to inquire whether the ultrasonic grunge emanating from his DAC3.1X zero-oversampling DAC, which I subsequently bought, would fry my T2 amp. It just shows that when we don't have the words to describe something accurately, we end up using whatever available existing descriptions, however erroneous.)

In the case of the RS-1, it is less a matter of coloration than it is of the headphone's rendition of mechanism #3. Headphones that render textures vividly sound more up-front. The language that audiophiles use in describing sound has become too dependent on descriptions of tonal balance. If a headphone is more up-front - blame it on the coloured tonal balance. If the headphone is more mid-hall, ascribe it also to the tonal balance. Everything becomes simplistically reduced to a matter of tonal balance. The effects of textural portrayal (#3) is not mentioned or not noticed.

Two tonally neutral headphones can sound different, despite their similar tonal neutrality. The headphone that renders #3 more vividly will sound more up-front and closer to the stage.

What challenges does #3 pose to the audio system?

Depth cue #3 requires that the audio system be capable of portraying textures vividly when the occasion calls for it, as well as portraying textures less vividly when another occasion calls for it. The challenge posed to the audio system is therefore textural range, to coin another new term. If "dynamic range" means the ability to portray the gamut of dynamics from fff to ppp, then textural range means the ability to portray the range of textures from less texturally specific to extremely texturally specific. *Textural range* means the ability to portray a highly textured sonic image alongside a not-so-highly textured image, such that a sense of depth is portrayed. It is not easy for audio systems to portray *textural range* accurately. Lesser playback systems tend to homogenize the sound, such that all textures tend to appear equally textured. Superior playback systems do not homogenize the sound, allowing textures of various instruments to come across as being texturally specific or texturally non-specific, independently of each other. *Textural range* is a key performance indicator of an audio system, especially in a headphone-based system where headphone-users have to rely

on comparative texture as a means of gauging spatial depth.

How well does the Omega II portray the Third Depth Cue?

Stupendously. The textures portrayed by this headphone can range from highly texturally specific to texturally non-specific, depending on what was in the recording. This headphone also does not homogenize sound, allowing a lot of breathing space for each texture to develop naturally and independently of each other. The textures of voices and instruments sound very different from album to album, which should be the case, as each album was recorded differently. And within the same album and same track, the textures also sound very different from one sonic image to another. Simply fantastic. Much of the spatial depth portrayed by the Omega II can be ascribed to its fantastic handling of *textural range*.

For example, I am now listening to the Track 20 (You Win Again) from The Very Best Of The Bee Gees. The insistent drum-beats sound distinctly further away due in large part to depth cue #3. (Drum-beats appear nearer when the textures of both the rattling drum frame and the taut dry drum skin being hit are abundantly present.) In the absence of both these textures, like for instance in this You Win Again track, the drum-beats seem to be further away, which is what I am hearing now via the Omega II. This headphone allows the texturally less specific drum-beat to co-exist with the more texturally specific voices. The texturally less specific sound of synthesizers creates the backdrop against which the texturally specific voices of the Gibb become the foreground object. I have found that rock music often employs synthesizers to create the backdrop against which foreground objects (typically voices) stand out. It has to do with the way synthesizers roll out smoother textures, and as #3 would have it, smoother textures sound more distant and can readily serve as soundstage backdrop. A handy little tool, the synthesizer.

Another example: Track 2 of Ali Farka Toure / Ry Cooder's Talking Timbuktu album. This CD is the collaboration between Ry Cooder who plays various sorts of electric guitars and Ali Farka Toure who sings and plays acoustic guitar and the njarka, accompanied by his team of Timbuktu percussionists. This album is filled with catchy melodies infused with pure and simple forms of rhythm. In Track 2, the percussive shaker is positioned dead centre of my forehead, but it sounds perceptibly distant. (The recorded sound of a percussive shaker placed up close to the microphone has a distinct texture, like the sound of many metal beads being agitated either by shaking or rubbing.) But in Track 2 of this album, the shaker definitely lacked such a high degree of textural specificity, implying the shaker's greater distance.

Another example: Track 5 (Amandral) from the same album includes a western drum kit, but the way it is played is deliberately subservient to the African percussive instruments during the opening and closing section of this track. The opening and closing of the track has the drum kit played such that the textural specificity of the air-squeezed double-cymbal and stick-hit cymbal is reduced. The reduced textural specificity of the air-squeezed double-cymbal and stick-hit cymbal contributes to their sense of greater distance, whilst the textures of the

calabash and congas remain highly texturally specific. This makes the western drum kit seem further away, and therefore compositionally subservient to the nearer-sounding African percussions. Then in the middle section of this track, the western drum kit acquires equal status to the African instruments. In this middle section, the leg-operated tambourine rips through the acoustic space with its clear vivid texture, appearing as forward sounding as the African percussions. Altered depth is used as a compositional element in this track, and this altered depth is achieved by altering textural specificities (#3).

Another example: 1st Movement of Shostakovich's Piano Concerto No.1 - the piano-trumpet duet sounds nearer to the listener than the accompanying orchestra. When the cello starts to play, I infer that it is further away because I hear neither the typical resinous purr of a string being bowed nor the typical woody resonance of a cello's body. Both the piano and trumpet are perceptibly more texturally specific than everything else, the piano more so than the trumpet. (It is after all a piano concerto.) The texture of the piano is highly specific - I am very aware of the percussive nature of the piano, its leading edge transients coming across sharp and clear. However, because the leading edge lacks the sharpest of bites, I also infer that I am not that close to the piano - I am not on the stage with the piano. I can understand the mental calculations involved in the recording engineer's mind when capturing this piece. On one hand, he must have wanted the piano to sound quite close because Shostakovich experiments here with "off-key" tonalities, and off-key tonalities on a piano sound best when captured near-field. On the other hand, he had to make the piano "gel" with the rest of the orchestra and cannot afford to have the piano stand out in too stark a relief against the accompanying orchestra. Hence the near-but-not-too-near perspective of this piano.

Strangely, as distance increases, different instruments lose their textural specificity at differing rates. For example, I am now listening to the 3rd and 4th movements of Beethoven's 5th Symphony - the part where sunshine bursts on stage when the brass section rejects the C Minor key in favour of the C Major. It is my observation that massed strings acquire a smooth texture whereas massed brass still retains a slight hint of the "brassy" texture. Maybe the higher harmonic textures of some instruments get attenuated faster than the textures of other instruments?

#### **DEPTH CUE #4:**

**sonic images swathed in a diffused/reverberative halo appear further, and this cue takes precedence over all other cues**

Hypothetical scenario: You are jungle trekking at night when you suddenly find a strange entrance in a stone cliff, covered by vines, into what you suspect might be a tunnel through the stone cliff. You adventurously go into the dark tunnel without any torchlight, relying only on your sense of touch and hearing to guide you. You have gone some 30 feet into the pitch-black tunnel (well I did say you were adventurous) when you suddenly realize you have passed from the tunnel into the belly of a large cave. Even in pitch darkness you knew you have progressed into a cave because you hear the fluttering of a thousand bat wings echoing

off the walls of the cave. The echoes of the fluttering wings “light up” the cave walls, and for that short duration when the echo could be heard you can “see” the extent of the cave walls.

Music is tied to architecture. I am not talking of the metaphorical relationship between music and architecture (that music is architecture in motion, or that architecture is frozen music). I am talking of the literal relationship between music and architecture - that some forms of music are so inextricably connected to the venue it is played. Choral and orchestral music is better heard in halls, and best heard in certain halls. Such music played in the open outdoors loses its usual sense of lushness.

Reverberation in recorded music occurs when sound is reflected off the walls, floor and ceiling of a recorded venue, and the microphones capture both the direct sound and the reflected sound that comes milliseconds after the direct sound. When you are nearer to the instrument, the amount of direct sound overwhelms the amount of reflected sound. When you are further away from the instrument, the ratio of reflected sound to direct sound gets larger. This gives rise to depth cue #4: whenever a sonic image is diffused with a reverberation halo, you perceive that that image is further away. I have consistently found by listening to recordings that depth cue #4 takes precedence over all the other three cues.

Depth cue #4 comes in two incarnations - overlapping reverberation (#4a) and impulse reverberation (#4b).

Overlapping reverberation (#4a) tends to occur with continuous sound sources, such as blown or bowed musical instruments as well as choir voices, whereas impulse reverberation (#4b) tends to occur with struck or plucked musical instruments.

Overlapping reverberation (#4a) is the reverberation that overlaps with the direct sound of a blown or bowed instrument whilst the instrument is still playing. The net result of this overlap is that the sonic image of the blown or bowed instrument acquires a certain “halo of diffusion”. Depending on the type of instrument and the hall characteristics, there might a core at the centre of the halo. Some diffused images do not have a central core; some do. I find that instruments that give off high-pitched textures tend to retain this core. Amazingly, sometimes the core can be so sharply delineated (because the core is texturally specific) that the core appears nearer (via depth cue #3) while the halo appears further. Curious.

(Because the overlap between direct sound and reflected sound causes a diffusion of the sonic image, I also call this type of reverberation “diffused reverberation”. Overlapping reverberation and diffused reverberation are one and the same thing.)

Impulse reverberation (#4b) is when the transient sound starts and then stops quite abruptly, with the reverberation quickly following in its wake. This occurs mainly with struck or plucked musical instruments. There may even be a very brief gap between the end of the direct sound and the start of the reverberation, similar to what you find in an echo. The reverberation also starts and stops quite abruptly, hence the name “impulse reverberation”. During the short duration of the impulse reverberation, the edges of the recorded venue

“lights up” momentarily but dramatically. Nothing, and I truly mean nothing, “lights up” the recorded venue quite as dramatically as impulse reverberation (#4b). It is as if you were a blind person but for a brief miraculous moment you were given the gift of sight. Quite wondrous really.

An example of impulse reverberation can be heard at the conclusion of the 4th movement of Beethoven’s 5th. The whole orchestra concludes in the C Major key in simultaneous syncopated bursts. Each burst is very brief, but very intense (because the whole orchestra contributes to the burst). A short moment after each burst, the hall “answers back” with an impulse reverberation burst, almost as if the reverberation note was on the composer’s score sheet. At those moments when the hall “answers back”, I can “see” the limits of the acoustic space.

Sometimes reverberation can be applied electronically, but I have found post-event reverberation to sound odd at times, and at rare occasions, truly hilarious. (The most comical application of electronically-added reverberation was in this particular piece where the female voice came from extreme left and the reverberation of her voice came from extreme right, and all through this piece there was a pretension of simulating a real acoustic space.) I find it acceptable to hear electronically-added reverberation if it was done in a witty manner or if there were valid compositional reasons. Certain music forms like rock, which is a form of amplified music, have no pretensions of being played in a natural acoustic setting, and if rock employs electronically-added reverberation I have often found that rather acceptable. The electronically-added reverberation was just one more electronic manipulation in a series of electronic manipulations like the judicious use of equalization and heavy mixing of multiple close-miked sources. I’m all right with it so long as there is no failed pretension at simulating a real acoustic space.

What challenges do depth cues #4a (diffused reverberation) and #4b (impulse reverberation) pose to the audio playback system?

The proper portrayal of #4a and #4b requires that the headphone playback system be (i) transparent such that there is little or no loss of ambient information contained in the recording, (ii) highly resolving such that each sonic image has ample breathing space and (iii) nimble-footed with quick transient response so that you perceive a heightened sense of real instruments playing in real acoustic environments.

How well does the Omega II portray depth cues #4a and #4b?

STAX headphones have a great tradition of being able to reproduce hall ambience excellently. There is an ethereal magical chemistry between STAX electrostatic headphones and reproduction of hall reverberation. STAX headphones have a light nimble touch that gives us the sense of real instruments hovering in real acoustic spaces.

The Omega II does not significantly depart from such pedigreed lineage. But the Omega II does not portray depth cue #4a (diffused reverberation) as vividly as other STAX

headphones like the Lambdas and the Omega I. The restrained upper-midrange and treble of the Omega II prevents the upper-midrange harmonics of ambient air from being “lit” brightly enough. There is no lack of transparency and resolution - via the Omega II you can hear right to the very rear of the soundstage, but it’s as if all the lights had been turned off and the recorded venue is plunged in darkness. The Omega II’s incredible transparency and resolution impart a very clear portrayal of ambient cues and hall reverberation. It is a superbly transparent window to the acoustic hall - it’s just that the Omega II is an utterly transparent window to a darkened hall, rather than a moderately transparent window to a more brightly-lit hall.

*Sidetrack: For this reason, I frequently turn off all the lights in my listening room when I listen to headphones - the actual darkness of my listening room complements the apparent darkness of the recorded venue. If I had a wish list for the new Omega III (if and when it comes out), it would be that the Omega III shines a little more light on the upper midrange spectrum of ambient air. Just a little more, but no more than that; or else the presentation would sound a little too “hi fi-ish”. It is a very tricky balance to get right.*

Other than this slight gripe, the Omega II is clearly superb in rendering hall reverberation and depth cue #4. For example, it is able to afford me an instructive demonstration of depth cue #4a (diffused reverberation) in Johann Strauss’s Explosions Polka 4th movement (Banditen Galop). The first explosion at 0.07sec seems reasonably nearby, while the second explosion at 0.11sec sounds further away than the first explosion because there is a greater reverberative diffusion (#4a) around the image of the second explosion. Coupled with this, there is also a sense of harmonic shift (#2b) with the second explosion that was absent in the first explosion. The third explosion at 0.19sec sounds even slightly further than the second explosion; this sense of greater distance was contributed by greater degrees of both #2b (harmonic shift) and #4a (diffused reverberation) relative to the second explosion. The location of the image of all three explosions remained the same: they were all located just beyond the left temple of my forehead.

#### **#2 + #3 + #4 + Air btw instruments:** **THE SENSE OF PERSPECTIVAL AIR**

Now I want to share with you something really magical called *perspectival air*. (There may be no such word as “perspectival”, but “*perspectival air*” is a newly coined phrase given definition and meaning in this essay.)

When two or more of the mechanisms combine, you get a greater effect of depth. Most convincing is when a single sonic image demonstrates #2, #3 and #4 simultaneously, coupled with a strong sense of air around the image. This combination of #2 + #3 + #4 + Air offers a devastating sense of *perspectival air* (played over the right headphones and set-up) - *perspectival air* to die for.

For example, I am now listening to Chris McGregor’s The Brotherhood Of Breath (a VTL

Recording using an all-Manley recording set-up). Pinise Saul sings into the mike (of course - how else would it have gotten into the recording?), but her voice is *not fed into the mix yet*. Her voice plays through a public address system, then the reproduced voice travels through 12-15 feet of air before being picked up by the main microphones. The acoustic 'haze' surrounding her voice is a joy to listen to, as is her singing. This 'haze' is achieved via mechanisms #2, #3 and #4, meaning to say that her voice sounds a little "tonally washed-off" (#2), loses quite a bit of textural specificity, for example the pronunciations of consonants are not as sharp compared to if her voice had been directly fed into the mix (#3) and the image of her voice is surrounded by a diffused halo of reverberation (#4). The combination of these 3 operative mechanisms plus the sense of air around the image of her voice gives rise to a tremendous sense of *perspectival air* - I am very much aware that the public address system from which her voice emanates is located some distance from the main pick-up mikes. Excellent stuff. *Perspectival air* to simply to die for.

Likewise, the plucked bass guitar in the same track is not fed directly into the mix, but played through the guitar speaker; the reproduced guitar sound then travels through intervening air before reaching the main mikes (the same main mikes that picked up her voice). This results in the bass guitar sounding airy, which may strike bass junkies as being odd - how can bass be airy? Bass is supposed to be solid and punchy, isn't it? Not really. (But more on this later.)

What is the difference between *perspectival air* and soundstage depth? After all, both occur in the z-axis (x-axis being left-to-right and y-axis being height).

Air may be the medium of transmission of sound, but air is also the *medium of resistance* to sound. The further sound travels through air, the more its volumetric (#1), tonal (#2), textural (#3) and reverberative (#4) character changes. *Perspectival air* is about the heightened aesthetic awareness that air is a medium of resistance to sound. The difference between "soundstage depth" and "*perspectival air*" is that the former is (merely) a perception of the z-axis, whilst the latter is about perceiving that the sound of instruments had to surmount an obstacle (air) in order to reach the microphones.

*Perspectival air* is a more acute and intense form of soundstage depth. You perceive soundstage depth when a sonic image displays any one or more of the Four Depth Cues. But when you get a potent combination of #2 + #3 + #4 + air around the instruments, you perceive glorious bountiful *perspectival air*. Without the fourth ingredient (air between the instruments) *perspectival air* will also be lacking. When only #2, #3 and #4 are present but the sense of air between instruments is lacking, what you get is soundstage depth, not *perspectival air*.

Most recordings give you soundstage depth, but not all recordings give you *perspectival air*. To give you *perspectival air*, the album has to be well recorded, most preferably minimally-miked, with ample ambient cues captured by the pick-up mikes. However, not all minimally-miked recordings give you *perspectival air* - production labels such as Clarity Recordings for

example offer a rather close perspective lacking in perspectival air despite their productions being minimally-miked.

Binaural recordings feature a lot of *perspectival air* by virtue of the minimalist approach of placing miniature microphones at the opening of the ear canals of a plastic dummy head. But I have yet to hear a binaural recording that gave me out-of-the-head imaging because I have yet to find a binaural recording that utilized a dummy head whose specifications exactly matches my personal HRTFs (Head Related Transfer Functions). But despite the usual in-the-head headstage that I experience with binaural recordings, such recordings gave me a headstage filled with a marvellous sense of *perspectival air*. No regrets there in having bought a total of 20-odd binaural CDs, even if I did not get the out-of-the-head experience that I thought I would get.

Labels such as VTL, Chesky, Mercury Presence, Telarc, Stereophile and Reference Recordings (amongst many others) feature recordings that have *perspectival air*. I have always thoroughly enjoyed the recordings released by such production labels when played over my headphones, but it surprised me to read at least 3 posts at Head-Fi that consistently complained about “the sense of distance” captured in such recordings. I cannot remember the threads or the persons who posted such a comment - but I was extremely perplexed by this consistency with which “sense of distance” automatically deserved criticism and rejection. Why would a headphone-user complain about recordings that portray depth cues or a lush sense of *perspectival air*? One answer might be that the audio system they own is not transparent enough to make sense of such recordings; another explanation might be that they have not yet acquired the experience to enjoy such recordings.

I have found STAX headphones to make me peculiarly aware of *perspectival air* - when it is present in the recording. I have owned five STAX headphones over the past 11 years (Gamma Pro, Sigma Pro, Lambda Signature, Omega I and Omega II), and can attest to the unique presentation style of STAX headphones. All the observations you read here in this essay have been slowly gathered by me over the past decade based on what I hear via those five STAX headphones, especially the Lambda, the Sigma and the Omegas. (The other headphone that presents an unsurpassed sense of *perspectival air* is the Sennheiser Orpheus.) I am not a recording engineer and I have not done recordings in my life before, nor am I a psychoacoustician, so it is highly curious that I can articulate several sonic phenomena that one would expect to be within the province of recording engineers or psycho-acousticians. This says something about the transparency of STAX headphones, which allows a home-user in the comfort of his listening chair to reconstruct the spatial characteristics of the recorded event.

*Sidetrack: This may also explain STAX's choice of calling their headphones “earspeakers”, because this term “earspeakers” more greatly carries a connotation of distanced air than the term “headphone”. However, I think that the deference to a loudspeaker-centric terminology may be unnecessary and potentially misleading, because a pair of loudspeakers creates an intervening distance between its “headstage” and the listener, whilst the effects of*

*perspectival air is about the intervening distance between musicians and the microphones. Seen from this angle, the fact that STAX headphones are prodigious portrayers of perspectival air should not make them deserve the epithet “earspeakers”. Perhaps by “earspeakers” STAX meant that their headphones co-opt the ear flap the way loudspeakers do, and not that STAX headphones are prodigious portrayers of perspectival air.*

How well does the Omega II fare compared to previous STAX models when it comes to portrayal of *perspectival air*?

I would describe Omega I's soundstage as being especially charged with the sense of *perspectival air* and that Omega II's soundstage, while not lacking in the portrayal of *perspectival air*, is not as super-charged. The slightly brighter middle-midrange and upper-midrange of the Omega I shines the light on the midrange spectrum of ambient air, making the sense of *perspectival air* super-charged, as if the air molecules above and around the musicians and between the musicians and the microphones were frenetic with vibration energy. (This occurs only if the correct recordings are played via Omega I - recordings that have a lot of *perspectival air*.) But what the first Omega lacked relative to the second is the sheer effortlessly relaxed clarity of its successor.

*(Summarizing the essay so far: Before going into my next section I just want to pause and take stock of what we've covered so far and what still lies ahead. We've covered the headstage, the Four Depth Cues and this incredibly lovely thing called perspectival air. I will now need to complete my review of the Omega II. I reviewed the Omega II using a review methodology structured on the Four Depth Cues, but an assessment of a headphone's depth portrayal is not enough - there are other things to evaluate. I will be touching lightly on six additional aspects: Background Blackness, Portrayal of Details, Bass, Midrange, Treble and System Matching. The reason why I am lightly touching on these aspects is because I do not wish to usurp the significance of the headphone review methodology based on the Four Depth Cues.)*

## **ADDITIONAL REVIEW ITEMS OF OMEGA II**

### **BLACK BACKGROUND**

All too often with lesser headphones, you become aware of the black background only when the music becomes less complex - the transition from the passage with many instruments to the passage with few instruments seem also to be accompanied by a transition from 'busy' background to a quieter background. With the Omega II, you never transit from busy background to quiet background - the background is always quiet and black, no matter how many instruments there are.

I believe that the Omega II's refined black background is due to its near-zero distortion. I have gotten so accustomed to the absence of distortion that I have become sensitised to it. After getting used to the Omega II, I suspect that there must be many types of insidious

distortions exhibited by other headphones. I am not talking about the obvious sort of distortion where the amplifier clips or something like that. I am talking about subtle forms of distortions. When such subtle distortions are at vanishing low levels, you get this incredibly velvety black background.

## **PORTRAYAL OF DETAILS**

The Omega II is a refined headphone. It portrays a lot of details - but it does not shove the details in your face. Rather, it is relaxed and casual about its rendition of detail. It's quite a paradoxical experience - there's oodles and oodles of details, yet the presentation seems very relaxed.

After having lived with this headphone for 4 years, I have come to the conclusion that its supremely natural and relaxed rendition of details is the result of 3 co-existing qualities:

- (i) ample dynamic headroom, such that there is no sign of stress and strain,
- (ii) ultra-high resolution, such that images are clearly distinguished from each other, and
- (iii) a velvety black background out from which images emerge effortlessly

## **BASS**

Can you believe that the history of STAX headphones had been primarily motivated by the search for true deep bass? Yet it seemed to be so. Years ago I read somewhere that in the mid-80s, when the Gammas used to be the top-of-the-line STAX headphones, the makers of Mercedes Benz cars needed a transducer that could tell them precisely what sort of low-frequency chassis resonance was happening in automobile frames. Thus was the first Lambda born - for a non-audiophile, non-recording industry purpose. Subsequently the Omega I appeared in 1992. The pamphlet for the Omega I says this: "large circular transducers...can effortlessly reproduce the lowest conceivable notes". Then the Omega II appeared in 1998 and further ups the ante on bass reproduction: "a new gold-plated electrode that attributes to increased bass response". Every new model had been primarily about further improving the bass reproduction.

I have a feeling that with the Omega II, STAX designers felt that they have finally cracked the nut on how to make a headphone go really deep. Back at HeadWize I called the Omega II "the heavy-weight bass champion of headphones", and I wasn't excluding dynamic headphones. (But please note I didn't say heavyweight bass-slam champion of headphones.)

There are 3 aspects to bass reproduction - bass slam, lower harmonics of voices/instruments and lower harmonics of ambient air. (But why do people keep thinking that there is only one aspect to bass performance, which is bass slam?) The Omega II excels in all three.

Bass slam - this headphone displays tremendous bass slam, when the recording calls for it. It is not a trade-off between weight and definition - the Omega II's bass slam is both weighty and tight. (But because of its restrained treble, the perception of bass slam via the Omega II

may not be as hard-hitting as compared to a brighter headphone. The sense of a hard-hitting drum is attributed more to the presence of high frequency textures than to low frequency weight alone.)

Lower harmonics of voices and instruments - this is even more important to me than bass slam because not all recordings call for bass slam but all recordings will benefit from a rich reproduction of lower harmonics. A deep, rich bass makes the tonal character of voices and instruments so much more authoritative and weighty. No headphone I've heard sounds as authoritative and weighty as this one.

Lower harmonics of ambient air - this is also very important to me, especially when I play albums that feature a lot of perspectival air or albums that feature harmonic shifts (depth cue #3b). No other headphone I've heard tells me so convincingly that hall reverberation also comprises of low frequency harmonics. People say that bass is matter of solidity, but I beg to differ. Bass to me is a matter of air as well. There is such a thing as a low-frequency ambient air - when you play large-scale orchestral works, it is the lower harmonics of hall reverberation that gives a sense of architectural scale to the music. The sense of weight and gravitas to music - this is Omega II territory.

## **MIDRANGE**

The all-important midrange, where most of the music is. Magical is how I would characterize the Omega II's midrange. I really dislike the phrase "smooth liquid midrange" because it is so overused, but I cannot think of a better phrase to describe the Omega II's midrange. There is nothing to dislike about the Omega II's midrange and everything to love.

Also, it is never just how this headphone portrays its midrange, but how the supporting bulwark of qualities such as velvety black background, ultra-high resolution and casual clarity come together to offer a clean, clear and sweet midrange.

One important thing to mention about the Omega II's midrange is that it is so fused with its treble and bass, that all the sonic images seem cut from the same cloth. The differentiation into bass, midrange and treble is in fact an artificial division. When you hear a trumpet via the Omega II, you don't just get midrange richness - you get the sound a trumpet that comprises the midrange principal harmonic plus upper harmonics plus lower harmonics all fused together to make the complete sound of a trumpet. "What midrange? I only hear a trumpet."

## **TREBLE**

The treble of the Omega II is difficult to describe. I have not read any review whether in HeadWize or Head-Fi or any professional magazine that accurately described the Omega II's beguiling treble (including my own review in 1999).

Quantity-wise, the treble of the Omega II errs very slightly on the side of insufficiency. Quality-wise, the treble of the Omega II packs oodles of clarity and resolution. Calling the

headphone “dark” is somewhat true, but only half the truth. “Dark” carries the connotation that the treble is soft-sounding, and this is true of this headphone to a certain extent. But “dark” also carries the connotation that the treble is muffled or not clear enough, and nothing could be further from the truth, for the Omega II is capable of resolving very finely textured treble detail. Its treble seems finer than silk - so fine that you can journey between the super-fine grains all the way down down down to the noise floor of your amp and source components.

This strange combination of a superbly fine-textured treble, yet shy treble, results in a headphone that is revealing-yet-forgiving. Because the treble is very finely textured, you can hear upstream nastiness like sibilance and smear, even in small amounts, but because the treble quantum is subdued, the upstream treble nastiness loses much of its sting, which accounts for the headphone’s forgiving nature. Revealing yet forgiving: the secret is in its treble.

This type of treble is a slight departure from absolute tonal neutrality. It errs on the side of warmth. But one good turn deserves another: I am willing to be forgiving of the Omega II’s tonal warmth, because it has been forgiving of my less-than-stellar recordings (of which I have plenty as well). Its revealing-yet-forgiving treble goes a long way in making my entire collection of CDs listenable and also in reducing listening fatigue to near-zero levels.

## **SYSTEM MATCHING**

Tricky issue to deal with. If you are a long-time owner of previous STAX models, you would welcome the Omega II’s non-fussy coupling with all sorts of source components and cables. This is because the Omega II does not sound as bright as previous STAX models such as the old Lambdas, which were more fussy about the tonality of system matching.

But if you are new to STAX headphones and you belong to the category of people who prefer up-front immediacy, then system matching becomes a more pertinent issue. When I first bought the Omega II, I was using the Muse Model 2 as my digital-analogue converter, which I would characterise as a little laid-back. I thoroughly enjoyed this partnership. (I’m a transparency freak, and I don’t really need up-front immediacy.) Then I bought the Audio Note DAC3.1X non-oversampling DAC. Audio Note DACs are musically lively, possibly due to the zero oversampling design, and it transformed the Omega II’s presentation into something more musically lively. I would say that the Omega II + Muse would not have appealed to people looking for greater immediacy, but Omega II + Audio Note - now that might rock your boat.

## **CONCLUSION (FOR REVIEW OF OMEGA II)**

The Omega II is a beguiling headphone. It has unique headstage characteristics (slightly frontal, small-sized, fulsome, hyper-focused). It portrays the Four Depth Cues well, in particular it has a most amazing *textural range* (#3), which greatly helps the listener in using

comparative textures as a means of gauging spatial depth. It portrays diffused reverberation (#4a) and impulse reverberation (#4b) well, with a sense of real instruments playing in real spaces, but the upper-midrange spectrum of hall ambience could do with a little more illumination. It portrays *perspectival air* (#2 + #3 + #4 + air) well, when it is present in recordings, although previous STAX models render *perspectival air* more vividly. It presents sonic images that emerge out from a quiet black background. It has an unbelievably prodigious yet tight bass, and it often portrays ambient air filled with low frequency harmonics, which imparts a sense of architectural scale to music. It has a magical see-through midrange that is uncannily cohesive with lower and upper ranges. It has a treble that is a little restrained but highly-resolved and refined. And the quality I cherish the most: it has a resolution and clarity so effortless as to become casual and relaxed.

The Omega II is a long distance runner. It is such a fatigue-free headphone that it can be used in an intensive manner by a compulsive headphone user (ahem!) who wears his headphone for a minimum of 4 hours at a single sitting, twice or three times a week, week after week, year after year (but with intermittent periods of complete rest, lasting 1-2 months each, to give the ears a necessary break and also to give myself a rest from too much of a good thing).

Is the Omega II the best headphone in the world? That's a very broad question, as there are many aspects to consider. But four aspects of the Omega II strike me as being possibly unsurpassed by any other headphone, dynamic or electrostatic. First is its clarity and resolution - no other headphone I've heard portrays such effortlessly casual and relaxed clarity. (There may be other headphones that match the clarity of the Omega II, but not its sense of relaxed clarity.) Second is its prodigious spectral weight - no other headphone I've heard sounds more authoritative and mature as the Omega II. Comparing all other headphones to the Omega II is like comparing the prepubescent voice of a boy to the voice of a matured man. Third, its midrange is so coherently integrated with the lower and upper reaches. Fourth, I have never heard a more finely textured treble from any other headphone.

So back to the earlier question: is the Omega II the best headphone in the world? My feelings now about this matter is: so what if it is and so what if it isn't? It is an irrelevant question for me now. This headphone has made me thoroughly enjoy a diverse range of music forms. It is as comfortable with classical as it is with rock (although I wouldn't describe it as a dedicated rocker's headphone that can play rock and only rock superlatively). It renders various forms of music with a great sense of ease and musicality and has kept me enthralled in this headphone hobby for 4 years (and running).

Talk about an extremely worthwhile investment.

### **CONCLUSION (A UNIQUE LANGUAGE FOR HEADPHONES)**

Listening via headphones offers a different realism from that offered by a pair of loudspeakers. A different reality requires a different language to describe it. A language that

specifically describes the sound of headphones has heretofore been either absent or under-developed. This essay seeks to fill that void.

The set of new words elaborated in this essay may be utilised to describe and review any headphone. The only reason I used this new language to describe and review the Omega II was merely one of convenience - the Omega II is after all my day-to-day headphone.

People who scoff at headphones for not portraying depth have not been listening alertly enough. While it is true that loudspeakers portray depth more convincingly, headphones DO portray depth, and they do so via four cues - volumetric (#1), tonal (#2), textural (#3) and reverberative (#4).

Granted, through a pair of loudspeakers you not only hear the Four Depth Cues, you can actually localize the externally located sonic images as well. In headphones, you do not have the benefit of externally located images, but you can train your ears to be more perceptive of distance cues inherent in recordings. Headphones are not deficient when it comes to portrayal of the Four Depth Cues, as I have been at pains to illustrate in this essay. (But headphones do lose out to loudspeakers when it comes to the One mechanism of sound localization.)

Come to think of it, the fact that the Four Depth Cues have been articulated as a *coherent paradigm* within the headphone world first and has not surfaced yet within the loudspeaker world suggests a possibility that headphones make us more aware of these depth cues than speakers do. Perhaps loudspeakers' localization ability is at once both an advantage and a handicap. If you have the *convenience* of externally-located images to give you the perception of depth, then would you be so acutely aware of the Four Depth Cues? Whereas a headphone-user who does not have the mechanism of localization at his disposal is *forced* to maximize his perception of the Four Depth Cues to grasp the spatial world of the recorded venue.

Will this essay be successful in instigating the growth of a language peculiar to headphones? I can only hope so.

May I politely request that Head-Fiers use some of the new words introduced here in their own posts and reviews? I have introduced many new words in this essay, but I wish to make the strongest case for only a few. *Headstage* is a word we cannot do without, once you understand what it means - what else are we headphone-users going to call that head-hugging soundfield that has kept faithful company? *Perspectival air* offers so much pleasure via headphones that it deserves to be used more often in order to describe those recordings or headphones that portray the sense of depth with such haunting airy realism. *Textural range* is a key performance indicator of a headphone's ability to portray depth via depth cue #3 - what other more appropriate word can we find to refer to that ability to portray spatial depth via comparative textures ranging from the non-specific to the highly specific? The term '*textural range*' is as appropriate and useful as the term 'dynamic range'.

There is really a chance here for the headphone community to craft a language peculiar to headphones. But someone has to first volunteer to produce the 'first cut' for everyone to debate and discuss. This essay is such a 'first cut'.

I've finally come to the end of this essay. Have a good day, everyone. I will be taking a long break after this exhausting write-up. Enjoy this wonderful little hobby of ours. Bye!

### **Footnote-essay no.1:**

#### **HOW TO PERCEIVE DEPTH CUES BONDED TO THE HEADSTAGE**

Play music via your headphones, and close your eyes. In your mind's eye, draw a rectangle, approx 8" wide and 5" tall, with the bottom of this rectangle resting on an interpolated line that connects both ears. You will find that all the sonic images portrayed by your headphone will "fit" into this abstract rectangle that you have just drawn in your mind's eye. This abstract rectangle is the headstage.

All the sonic images are resting on this abstract vertical rectangle. ("Resting" is a strange word to use when music is dynamic.) Think of the sonic image as a child's sticker book sticker - in your mind's eye you paste this "sticker" on the flat rectangle. Sometimes the "stickers" may overlap each other, but don't be too bothered about this - it is natural for two or more sonic images to sometimes occupy the same space. If you own high-end equipment, it becomes increasingly difficult to picture the sonic images as flat "stickers" because the images seem so full-bodied and rounded to you. In which case, do not fret - think of the headstage as the vertical plane that intersects through the centres of all those full-bodied "balls of sound". Or think of the headstage as an upright rectangular tupperware that contains these rounded sonic images.

Concentrate on one sonic image. Precisely where on the rectangle is it located? Is it located nearer to the right edge of the rectangle? Is it located nearer the top edge or bottom edge of the rectangle? On lesser playback systems, it can become difficult to pin-point the precise location of the sonic image - the image seems to be smeared over a larger area. On superior playback systems the image location is precise and can be effortlessly located. Once you have determined the location of this image on the rectangle, you can proceed to the next stage. Of this sonic image you picked, ask yourself: is it soft-sounding? Then go the next question: is the image you picked tonally washed-out? Then: is it texturally washed-out? Then: is it swathed with a reverberate halo?

When you have run through all four mechanisms for the first sonic image, proceed to the next sonic image of your choice. Run it through the same checklist of five questions (its location on the rectangle and the subsequent four questions). When you are done with the second image, proceed to the third.

It all sounds very tedious, but it isn't. It is actually simpler than it appears in this write-up. (Either that, or I've had a lot of practice.) It isn't really a chore because you have to

remember: you are bobbing your head up and down to the rhythm and melody of your favourite music. (Either that, or you're waving your imaginary baton in empty air.) How can that be a chore? If anything, the awareness of each image's portrayal of the mechanisms only serves to deepen the enjoyment of music.

After some practice, the awareness of the planarity of the headstage and the perception of the Four Depth Cues come quite naturally. With practice the enjoyment of the music is integrated with the perception of depth cues. It seems counter-intuitive - the idea that in order to hear depth cues better you need to first focus on the planarity of the headstage plane. But keep practising at perceiving the planarity of the headstage and its Four Depth Cues and you will become a more discerning headphone listener who can quickly and accurately decipher the depth cues inherent in recorded music.

### **Footnote-essay no.2:**

#### **NON-CARTESIAN CO-ORDINATES**

If I were asked to paraphrase the headstage and its 4 depth cues into a computer programme code for processing of depth cues via headphones, I would create the following 8 variables:

$$(x, y, z, r) + (a, b, c, d)$$

where

x = left-to-right location of image

y = up-down location of image

z = 0, which will create a flattened headstage

r = radius or roundness of images

a = loudness of image (depth cue #1)

b = tonal richness (depth cue #2)

c = textural specificity (depth cue #3)

d = reverberation amount (depth cue #4)

You might notice that (x, y, z, r) are variables that arise out of the One mechanism of sound localization. And (a, b, c, d) are variables that each arise out of the Four Depth Cues.

Assigning z = 0 will create a flattened headstage. Variable x is simply about stereo panning and should be easy to programme for a pair of stereo headphones. Variable y is difficult to programme - what gives rise to the sense of up and down placement of images? Variable r is difficult to programme - what gives rise to a sense of roundness of images? Variable a is easy to programme - it is simply a matter of volume control. Variable b is simple to programme - it is simply a matter of equalization. Variable c is difficult to programme - how does a computer programme increase and decrease the "trumpetness" of a trumpet? A computer cannot recognize the texture of a trumpet simply from wave analysis. Variable d is

simple to programme - it is a matter of feeding slight delays to the original sound. But using a computer programme to simulate good hall ambience must surely be an art form.

### **Footnote-essay no.3:**

#### **CAN YOU INCREASE HEADSTAGE SIZE?**

To increase the headstage size means to create images that are located further from the head, even to the point of creating out-of-the-head images.

The only way to significantly enlarge the headstage is to listen to binaural recordings, but as I've noted previously, it's unlikely for the your personal HRTFs to coincide with the dummy head used in the recording. Consequently, most of us will still experience an in-the-head headstage when listening to binaural recordings.

But there are some options open to you if you wish to slightly increase the headstage size. (Keyword = slightly.)

The headstage is the result of the transducer's location in relation to your ears. I have not auditioned them before, but I would imagine that Jecklin Float headphones create slightly larger headstages than most other headphones, simply because the left and right transducers in a Jecklin Float (and AKG K1000 as well, come to think of it) are about 2 inches wider apart than almost all headphones. This increased distance should create a slightly larger left-to-right soundfield, i.e., a wider headstage, but I'm not speaking from firsthand experience of the Jecklin Floats here. Swivelling the K1000's earpieces frontally should create a most amazingly frontally-located headstage, unrivalled by any other headphone probably except the STAX Sigmas.

The tonal character of a headphone has a small but perceptible effect on headstage width and headstage height. Brightness in the middle-midrange and upper-midrange results in slightly taller headstage heights when playing distance-miked recordings, but results in a solidifying of sonic images when playing close-miked recordings with no apparent effect on headstage size. Brightness in the upper treble has the effect of slightly increasing the headstage width in close-miked recordings, but slightly increasing the headstage height in minimally-miked recordings. I am generalizing here - not all close-miked recordings sound the same and not all minimally-miked recordings sound the same. But my central point here remains valid: the tonality of a headphone or recording slightly affects the resultant headstage size, either in width or height or both. I must emphasize the 'slightly' part.

But a larger headstage is not necessarily better than a smaller headstage. It's bit like saying that a 6" photo is better than a 5" photo. Is it really? Of course it is nice to have a larger headstage (in the same way it is nice to have a larger computer monitor), but how about comparisons between clarity, resolution, texture and colour saturation? Size of headstage is only one consideration out of many. Moreover, the differences between headstage sizes of various headphones are not that significant (at least from my experience), so it really

becomes less important to compare headstage sizes. If the difference between the 5" photo and the 6" photo is only 1" (if my math isn't rusty), and the 5" photo has better clarity and colour saturation, then why not go for the 5" because more significant factors outweigh the small gain in size?

#### **Footnote-essay no.4:**

#### **WHAT DOES "LARGE SOUNDSTAGE" MEAN?**

Now and then I come across posts at Head-Fi that says headphone X when coupled with headphone-amp Y creates a "large soundstage". What does the phrase "large soundstage" mean in perceptual terms? In abstract terms we all know what large soundstage means - it means that the soundstage is large. Duh. But what exactly did the person perceive that prompted him to use the term "large soundstage"?

It could be any one of five possibilities that prompted him to use the term "large soundstage":

(i) the headstage itself has increased in size, i.e., meaning that a sonic image instead of being located in its usual position touching the left temple (for instance) has now suddenly acquired through happy accident an illusion of being located 3 inches in front of the left temple (for instance). By happy accident I mean a freak coincidence where your personal HRTFs and the phase/frequency peculiarities of the recording and phase/frequency peculiarities of a headphone system commingle to result in a binaural-like illusion of an externally-located image hovering 3 inches beyond the left temple. This occurs very rarely. It is extremely rare for images to drift away to some out-of-the-head location.

(ii) the recording he heard had a deep soundstage and he could hear to the very rear of the soundstage. The backdrop of the soundstage is created by a sonic image that portrays a depth cue or a combination of depth cues, and the depth of the backdrop is further emphasized by the presence of a foreground object. The foreground object is tonally richer or texturally richer or reverberatively poorer than the backdrop image. Experiencing this clear background-foreground relationship may account for another possibility why a person would say he hears a "large soundstage".

(iii) the recording had ample reverberation cues and his headphone is transparent enough to render such cues. Reverberation diffuses a sonic image and makes the sonic image acquire a halo around the image. The presence of this halo of diffusion results in a perception that the sonic image has increased in size as well as making the image sound further away (#4). The bigger and subjectively further sonic image leads to a subjective perception that the soundstage has correspondingly increased both in lateral size and depth.

(iv) strangely, some instruments tend to "stand tall" in the acoustic space. Choir voices and horns tend to do that. I have no idea how or why this occurs. The Four Depth Cues only work in the z-axis, and I have not been able to account for mechanisms that work in the y-axis. So another possible reason for a person to say that he hears a "large soundstage" is because

he hears an image standing tall in the acoustic space, which contributes to his illusion of a larger soundstage. (v) smeared sound is mistaken to be a wide backdrop or a wide sonic image. This smear may be inherent in the recording or may have been introduced by the audio component.

### **A LEXICON OF HEADPHONESPEAK**

Headscape: the contours and landmarks of your head that you can reference the location of sonic images against

Headstage: the head-hugging soundfield resulting from the One mechanism of sound localization

Soundstage: perception of width via left-to-right differentiation, perception of depth via Four Depth Cues and perception of height via an unknown mechanism

The Four Depth Cues: the four mechanisms based on the principle that the further sound travels through air, the more its volumetric (#1), tonal (#2), textural (#3) and reverberative (#4) character changes.

Depth cue #1: soft-sounding images tend to appear further.

Depth cue #2: Tonally attenuated images tend to appear further. Occurs in two incarnations - tonal blandness and harmonic shift.

#2a: tonal blandness. Both higher harmonics and lower harmonics are attenuated, causing tonal blandness that makes an image appear further.

#2b: harmonic shift. Only the higher harmonics are attenuated but the lower harmonics are intact, causing an image to sound deeper and appear further away.

Depth cue #3: Images with reduced textural specificity appear further.

Depth cue #4: Images surrounded by a halo of reverberative diffusion appear further. Comes in two incarnations - diffused reverberation and impulse reverberation.

#4a: diffused reverberation. Reverberation that overlaps with original sound. Occurs with bowed and blown instruments as well as voices.

#4b: impulse reverberation. Reverberation that starts and stops quickly in the wake of the original sound's demise. Occurs with struck or plucked instruments.

Textural range: an audio system's ability to portray differing depths by portraying the whole gamut of differing textures

Perspectival air: a more acute (and more enjoyable) form of soundstage depth when mechanisms #2, #3, #4 and the sense of air around instruments comes together to form a heady mix