LINKWITZ LAB

Sensible Recording and Rendering of Acoustic Scenes

----- Experience in your own room the magical nature of stereo sound -----

... but you want to design your own loudspeaker?

My conversations with Fitz

What's new

LX - Store

Conversations with Fitz

OPLUG Forum

Basics

The Magic in 2-Channel Sound

2 - An introduction to Acoustics and Hearing

2.1 - Hearing - Part 1

SL: The Zen master asks the student: "When a tree falls in the forest, does it make a sound?"

And the student thinks: "That must be one of his trick questions again, I must think before I answer. I know when the tree falls that it makes noise when its branches hit other trees and branches and that it makes a loud crash when it hits bottom. so why is he asking?"

Fitz: I get it! Somebody has to be there to hear it. Otherwise it's just air vibration. Sound is what I hear. What I perceive and then associate with branches moving against other branches.

2.1 - Hearing - Part 1

SL: The Zen master asks the student: "When a tree falls in the forest, does it make a sound?"

And the student thinks: "That must be one of his trick questions again, I must think before I answer. I know when the tree falls that it makes noise when its branches hit other trees and branches and that it makes a loud crash when it hits bottom. so why is he asking?"

Fitz: I get it! Somebody has to be there to hear it. Otherwise it's just air vibration. Sound is what I hear. What I perceive and then associate with branches moving against other branches.

SL: The Zen master also asks: "What is the sound of one hand clapping?"

Fitz: That does not make any sound. But I can hear clapping in my mind. So the sound of one hand clapping is referring to a memory.

SL: Sounds reasonable to me, but then I am not a Zen master to know why he asked the questions. I suspect he is inquiring about the student's Self with a capital S, about the person who is hearing or not hearing, and what their state of enlightenment is at this moment. That would jibe with another master's response to a student's question: "What is the essence of Buddhism?" And he answers: "No Self, no problem!"

Fitz: That gets heavy. Let's talk about hearing.

SL: Yes, read first what Bregman had to say. His book was published 1990, not that long ago. He explores the amazingly complex processes, which the brain is capable of in order to make useful sense of what pressure variations in the air communicate about the world around us. It's a thick book and I did not make it much beyond the introduction, but it influenced tremendously my thinking about sound. Ultimately, and in combination with Guenther Theile's Association Model and my own experimentation, I learned that a loudspeaker must be friendly with natural hearing processes in order to render sound convincingly.

From: Albert S, Bregman,
Auditory Scene Analysis
The perceptual organization of sound

Sound is a pattern of pressure waves moving through the air, each sound-producing event creating its own wave pattern. The human brain recognizes these patterns as indicative of the events that give rise to them: a car going by, a violin playing, a woman speaking, and so on. Unfortunately, by the time the sound has reached the ear, the wave patterns arising from the individual events have been added together in the air so that the pressure wave that reaches the eardrum is the sum of the pressure patterns coming from the individual events. This summed pressure wave need not resemble the wave patterns of the individual sounds.

As listeners, we are not interested in this summed

To deal with this scene analysis problem, the first thing the brain does is to analyze the incoming array of sound into a large number of frequency components. But this does not solve the problem; it only changes it. Now the problem is this: how much energy from each of the frequency components, present at a given moment, has arisen from a particular source of sound, such as the voice of a particular person continuing over time? Only by solving this problem can the identity of the signals be recognized.

For example, particular talkers can be recognized, in part, by the frequency composition of their voices. However, there are many more frequencies arriving at the ear than just the ones coming from a single voice. Unless the spectrum of the voice can be isolated from the rest of the spectrum, the voice cannot be recognized. Furthermore, the recognition of what it is saying – its

From: Albert S, Bregman,
Auditory Scene Analysis
The perceptual organization of sound

Sound is a pattern of pressure waves moving through the air, each sound-producing event creating its own wave pattern. The human brain recognizes these patterns as indicative of the events that give rise to them: a car going by, a violin playing, a woman speaking, and so on. Unfortunately, by the time the sound has reached the ear, the wave patterns arising from the individual events have been added together in the air so that the pressure wave that reaches the eardrum is the sum of the pressure patterns coming from the individual events. This summed pressure wave need not resemble the wave patterns of the individual sounds.

As listeners, we are not interested in this summed pattern, but in the individual wave patterns arising from the separate events. Therefore our brains have to solve the problem of creating separate descriptions of the individual happenings, but it doesn't even know, at the outset, how many sounds there are, never mind what their wave patterns are; so the discovery of the number and nature of the sound sources is analogous to the following mathematical problem: "The number 837 is the sum of an unknown number of other numbers; what are they? There is a unique answer."

To deal with this scene analysis problem, the first thing the brain does is to analyze the incoming array of sound into a large number of frequency components. But this does not solve the problem; it only changes it. Now the problem is this: how much energy from each of the frequency components, present at a given moment, has arisen from a particular source of sound, such as the voice of a particular person continuing over time? Only by solving this problem can the identity of the signals be recognized.

For example, particular talkers can be recognized, in part, by the frequency composition of their voices. However, there are many more frequencies arriving at the ear than just the ones coming from a single voice. Unless the spectrum of the voice can be isolated from the rest of the spectrum, the voice cannot be recognized. Furthermore, the recognition of what it is saying – its linguistic message – depends on the sequence of sounds coming from that voice over time. But when two people are talking in the same room, a large set of acoustic components will be generated. These have to be stitched together in the right way. Otherwise illusory syllables could be perceived by grouping components derived from both voices into a single stream of sound.

From: Albert S. Bregman and Pierre A. Ahad, Demonstrations of Auditory Scene Analysis - The perceptual organization of sound, Audio CD

SL: And now another of Bregman's illustrations of the ear-brain system's capability. I like to call that system our BSP, our biological signal processor. Imagine writing the software for a DSP box that must function like the BSP. If the software does not work, you will be eaten by a saber toothed tiger.

Fitz: It would seem that 'room correction' could start a fight with the BSP.

SL: That is possible. I judge any sound system by how tiring it is to listen to. When after a while you feel like you had enough, then your brain is telling you subconsciously: I have worked enough to compensate for the unnatural cues that I receive. I am tired.

The difficulties that are involved in the scene analysis processes in audition often escape our notice. This example can make them more obvious. Imagine that you are on the edge of a lake and a friend challenges you to play a game. The game is this: Your friend digs two narrow channels up from the side of the lake. Each is a few feet long and a few inches wide and they are spaced a few feet apart. Halfway up each one, your friend stretches a handkerchief and fastens it to the sides of the channel. As waves reach the side of the lake they travel up the channels and cause the two handkerchiefs to go into motion. You

are allowed to look only at the handkerchiefs and from their motions to answer a series of questions: How many boats are there on the lake and where are they? Which is the most powerful one? Which one is closer? Is the wind blowing? Has any large object been dropped suddenly into the lake?

Solving this problem seems impossible, but it is a strict analogy to the problem faced by our auditory systems. The lake represents the lake of air that surrounds us. The two channels are our two ear canals, and the handkerchiefs are our ear drums. The only information that

Solving this problem seems impossible, but it is a strict analogy to the problem faced by our auditory systems. The lake represents the lake of air that surrounds us. The two channels are our two ear canals, and the handkerchiefs are our ear drums. The only information that the auditory system has available to it, or ever will have, is the vibrations of these two ear drums. Yet it seems to be able to answer questions very like the ones that were asked by the side of the lake: How many people are talking? Which one is louder, or closer? Is there a machine humming in the background? We are not surprised when our sense of hearing succeeds in answering these questions any more than we are when our eye, looking at the handkerchiefs, fails.

The difficulty in the examples of the lake, the infant, the sequence of letters, and the block drawings is that the evidence arising from each distinct physical cause in the environment is compounded with the effects of the other ones when it reaches the sense organ. If correct perceptual representations of the world are to be formed, the evidence must be partitioned appropriately.

In vision, you can describe the problem of scene analysis in terms of the correct grouping of regions. Most people know that the retina of the eye acts something like a sensitive photographic film and that it records, in the form of neural impulses, the "image" that has been written onto it by the light. This image has regions. Therefore, it is possible to imagine some process that groups them. But what about the sense of hearing? What are the basic parts that must be grouped to make a sound?

From: Albert S. Bregman, Auditory Scene Analysis - The perceptual organization of sound, MIT Press, 1990

Fitz: That is truly amazing.
... and what is this Theile Model all about?

SL: Well, I have to first generate some visuals to illustrate and explain the Association Model of auditory perception. We will talk about the philosophical concept of Gestalt and how we experience it. And then we talk about the ear and upper body as an encoding and decoding device for the world around us in order that the brain can extract meaning and learning in conjunction with its memory part.

It is not obvious at this time, but it all relates to loudspeakers and what they must do for a listener to recognize a Gestalt, and how therefore they should be designed and built optimally.

...Interlude...

2.2 - The Association Model of Perception

SL: I am back.

. . . I n t e r l u d e . . .

2.2 - The Association Model of Perception

SL: I am back.

I encountered the association model for the first time in Guenther Theile's AES Journal paper, "**On the Standardization of the Frequency Response of High-Quality Studio Headphones**", Vol. 34, No. 12, 1986 December.

I noticed the model again, and this time applied as a guide to effective coding of information, in a white Paper by Clemens Par.

You might read the papers, if you want more background material for our conversation about loudspeaker design and about how and what we hear:

Audio, however, is a highly controversial subject, as models how human hearing occurs are closely linked to cerebral activities, which are only investigated up to a specific degree. The foremost model, conceived as a doctoral thesis in the eighties, is the so-called "Assoziationsmodell" by my friend Günther Theile [4]. See Figure 1.

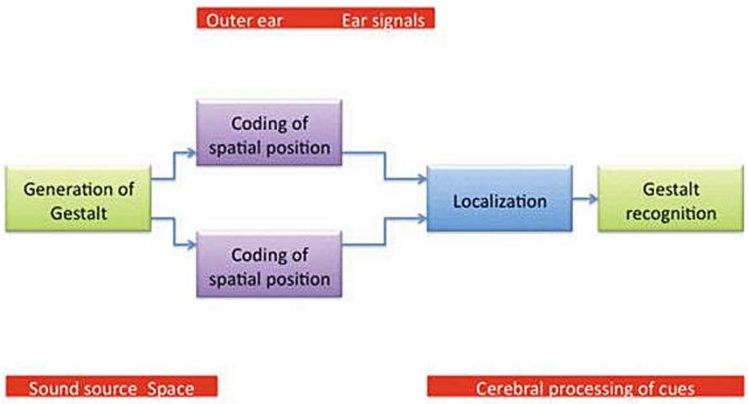


Fig. 1: "Assoziationsmodell" according to Günther Theile, describing the human auditory response to an external stimulus [4].

From: Clemens Par, Rationalism versus Empirism, in http://www.intercomms.net/issue-25/va-1.html, 2015
or http://www.linkwitzlab.com/Fitz/rationalism-empirism.htm
[4] G. Theile, On the Localization in the Superimposed Soundfield, PhD Thesis, Technische Universitaet Berlin, 1980

Rationalism versus Empirism

swissaudec® Sound science

A Crash Course in Invariant Theory and a Tribute to Rudolf E. Kálmán by Ecma TC32-TG22's Convenor and Swissaudec's CEO Clemens Par

Download article as PDF

Here is my friend Clemens looking to convince the engineering world to apply sensible perceptual coding

Fitz: I will skip it for now and maybe get back to reading the papers later.

2.3 - Perception of the world around us

2.3 - Perception of the world around us

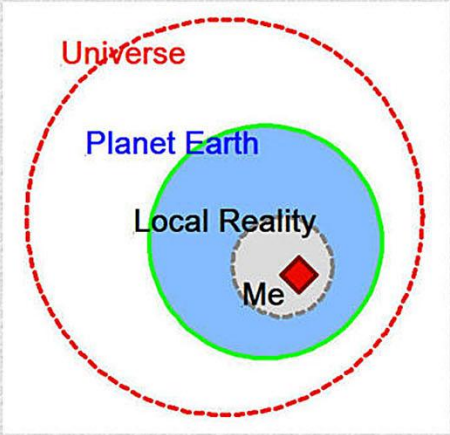
SL: Alright, but I must tell you that the association model gives not just a wonderful explanation to how we hear. It applies to all our senses, the five senses through which we experience the world on the outside of our skin: Hearing - Seeing - Touching - Smelling - Tasting.

Fitz: How is that?

SL: I really must step back a long way, because this is about how we interact with and respond to the world around us. It is about the functional patterns that we have acquired by birth from our parents and by generations before us, through the genetic code in our bodies. It is about learning, since the first scream at our birth into this universe, about experiencing love and pain and then learning to get more love and how to avoid pain. It is about the influence of 'nature' and 'nurture' upon our daily activities and longings.

Fitz: You are going big again.

SL: It hardly gets bigger than that. So lets start right away with the Universe.



Fitz: Yes, I know that Planet Earth and our solar system is just one of millions of solar systems in the universe. I read that they found just recently a cluster of planets, which potentially could carry life as we understand it, and the planets are not that far away in terms of light years.

I know that we are not the center of the universe, not even the center of our milky way galaxy, but some 20,000 light years away from the center and at the outer fringes of our spinning galaxy, which is just one of millions of galaxies. It is just overwhelming to think about it, but also awe inspiring and humbling.

SL: Yes, and then to think that our planet is just a youngster in the universe, recycled material from extinguished suns, from stars, which shone brightly before the earth was born. We ourselves, the atoms, which make up our bodies were cooked in a sun. We are literally made up of stardust. We are local condensations and manifestations of a parent sun.

Fitz: Pretty amazing stuff

SL: And the beauty is that we are, or can become, conscious of this relationship through the extension of our five senses by scientific instruments, and by communication tools that reach beyond the range of our voice or beyond the drum beats, which were in earlier times used to send messages to people at great distance from each other.

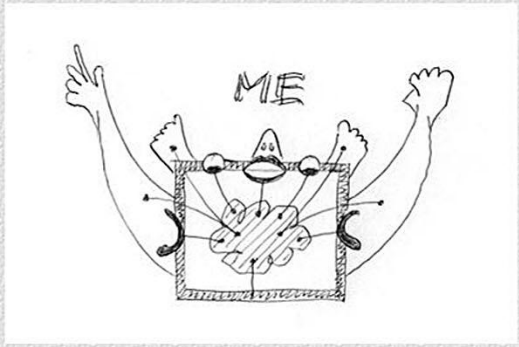
Fitz: OK. The picture above shows Me inside a Local Reality, which is inside the circle called Planet Earth.

SL: What is being communicated here, is that we live our lives locally. In the past the extend of that region was defined by the horizon on the water, or the next mountain range, or the forest, or basically by how far we could see. Compared to that, the local reality for hearing, what you can perceive in terms of spatial information with closed eyes, is different in its extension and detail. Touch is limited by the reach of your arms or legs.

Fitz: But today my Local Realty is far greater, because I have a smart-phone and ear-buds and a car and Facebook and Google.

SL: Yes, very true. But I still call it Local Reality, because you have self imposed boundaries, which determine how far you reach. You have a local world to which you respond. I have another one. Since you are on my website, our worlds have some amount of overlap.

The point of all of this, is that we deal with, that we experience and live in a bounded world. We can transport ourselves by foot, bicycle, car or airplane to different bounded worlds. Taken together, all these add to and form our experience and learning of the world beyond our skin. We carry in our head and body memories, thoughts, emotions, behavior patterns and models as a result of our journey through life. The journey started with birth. We all came into this world pretty much the same way. We go out in our own unique way. Between those two events we live the story of our life.



SL: How the story is being written depends upon the inputs and messages, which we have received in the past, and what and how we have internalized them. The story depends upon the inputs, which we receive at every moment and how we respond to them. But the story also develops in response to our resistances, expectations and hopes.

The difficulty for our perceptual apparatus, for hearing and seeing for example, is that we are constantly exposed to information from our local reality.

When I open my eyes in the morning, I see the reflections of the sunlight, which describe the objects in my bedroom and the objects outside the bedroom window. Since I have memory, which I associate with what I see now, I am not so much interested in the scenery and landscape, but I am looking for what has changed since yesterday. For example I like to know what today's weather is going to be. I know from experience that last night's weather report is not a reliable predictor here, on the Pacific coast of Northern California, where the fog drifts in and out, and the weather can change from sunny to foggy within an hour or two. So the weather has my attention.

Now if I stand up to look out the window, I notice whether the long grasses or any tree branches move, and if so, in which direction. Having been an avid, a practically addicted windsurfer, I cannot help but noticing the smallest movements caused by air, and hopefully indicating a windy day. I no longer dare to go out on the water, aging has its consequences, but I still respond to patterns, which I adopted in earlier times. Water, wind and waves are still calling, when my pleasure now is to wade through the cold waters along a sandy beach near our home.

Fitz: OK, why are you telling me this?

SL: I am telling you this because our perceptual apparatus has to divide the world we live in into foreground and background. Otherwise we cannot handle the inflow of information. We need to know what is background, what is consistent, what does not require attention, in order to notice changes in the background, which might carry information, to which we better pay attention. It is the movement against a background, whether it is a visual or a sonic background, which draws our attention.

As a matter of fact, all our senses are primarily change detectors. Come into a room and it smells bad. After a while you no longer notice it. But then go out the door and you will relish the fresh air.

Fitz: It's the same with hearing. I get really annoyed with clicks and ticks on my old LP's and often wish I could just ignore them.

SL: That is difficult to do. You can much more easily ignore tape hiss or LP groove noise, because it is continuous, Your brain can move that beyond its acoustic horizon. But is has to constantly work at it and it becomes tiring. Some audiophiles like to talk about new equipment brake-in. I think it is mostly their brain breaking in, because the brain has plasticity, and is eventually adapting to the reality. The longer the break-in time, the more artificial the sound must have been.

The clicks and ticks though draw your attention, because you are genetically programmed to pay attention to such signals. In evolutionary terms, there could be a saber toothed tiger in the bushes. We still respond to the commands from a large region in our brain, which is programmed for bodily response and motion, like hiding or running.

SL: Our conversation now brings me to the Association Model of Perception, or AMP. The general model came to mind as I was thinking about the steps and processes in the Association Model of Hearing, which GuentherTheile and then Clemens Par used for their particular lines of investigation.

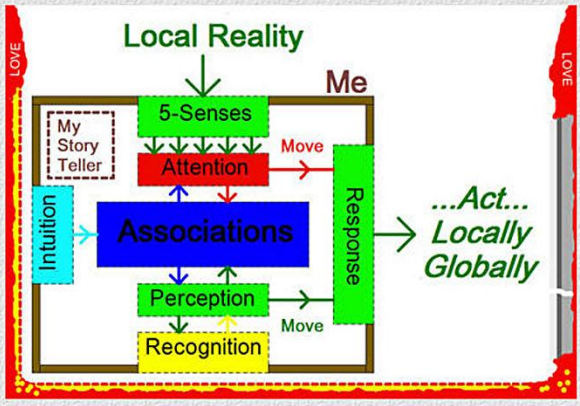
I also have been fascinated by books like:

Cesar Hidalgo, "Why Information Grows", 2015
Daniel Kahneman, "Thinking, Fast and Slow", 2011
Michael S. Gazzaniga, "Who's in Charge?", 2011
Pierre Teilhard de Chardin, "The Phenomenon of Man", 1955

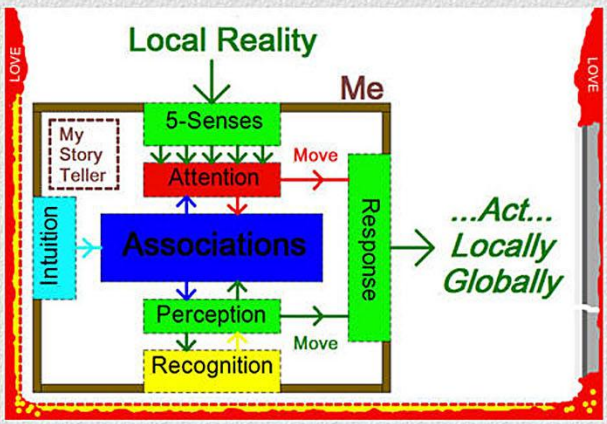
The model below brings together what I have taken from various inputs, from my own life's story and from recognizing what is hidden behind "The Open Secret" of life and living.

Fitz: So you have tested this model against your experiences and understandings?

SL: Yes. The beauty of this model is that it not only brings to consciousness the patterns in my response to the realities beyond my skin, but also to a dimension beyond, which touches and enters me through Intuition and Recognition.



Association Model of Perception



Association Model of Perception

Attention }-- Movement --{ Perception

SL: During Intermission it occurred to me that the **Association Model of Perception** is as important to understanding life and living, as is Einstein's ($e = mc^2$) to understanding physics.

The **AMP** describes how we take in the outside Reality and deduce from our perspective a Subjective Reality, which has Meaning to us in context of previous observations of Reality. We associate current input data with stored memory data, many of which are hard to access directly, and we respond accordingly: physically, mentally, emotionally and intuitively.

2.4

... Perception and the Perceiver ...

SL: I perceive my physical environment via five sensory transducers: eyes, ears, skin, tongue and nose. Electrical signals from the sensors are transmitted to the brain, which then deduces from them a model of the physical environment. The brain takes into account associations with memories, emotions, love, fear, traumas, pain, suffering, death, love, acceptance, laughter, joy, feelings, thought forms, beliefs, knowledge, education, patterns of thought, skills, convictions, experiences, etc, etc.. This can be a time consuming process before perception and cognition occur.

Survival of the species requires fast processing of incoming data, which is done by sorting the incoming data streams into static data streams and transient data streams. Transient or changing data streams get first attention. The external visual or acoustic scene is divided into background and foreground. The foreground gets attention and a response, which is most appropriate in the context of the present background. Decisions have been made and actions taken before I am even aware of what has happened. This occurs in the part of the brain, which is hard-wired for survival and with which I was born.

Perception and recognition, which occur after the input data has been processed via associations, leads to completely subjective responses. They are often not appropriate in the context of the external reality. They are instead responses, which come out of various forms of resistance to the external reality.

The Story Teller in your head is the commentator on your journey through life, which may take you from Self-consciousness to Consciousness of Relationships and Connectedness, to Enlightenment, and/or through Grace to Recognition. Along the way your story teller is ready to come up with excuses, justifications and explanations for all your actions and inactions.

The Story Teller in your head is the commentator on your journey through life, which may take you from Self-consciousness to Consciousness of Relationships and Connectedness, to Enlightenment, and/or through Grace to Recognition. Along the way your story teller is ready to come up with excuses, justifications and explanations for all your actions and inactions.

Regardless of the nature of the actions, and the consciousness with which they are done, they take place in an all encompassing energetic field of LOVE. That is difficult to accept, but becomes apparent when resistance is dropped.

SL: The song "Celestial Echo" from the Boris Blank and Malia album "Convergence" is about a story teller. Which story do you tell yourself?

Fitz: I want to go back to the two modes of mental operation. The first mode comes from my instinctive/intuitive side. The second one is from my subjective perception of reality, after the input data have been filtered by associations and maybe some intuition.

SL: Yes, and the whole journey through life is one of paying attention to inputs and to respond physically, mentally, intuitively and emotionally. What needs to be learned, is that resistance is futile. Battles are fought for Peace, not for Domination. Ours is a planet of Balance.

Fitz: And what about AMP and
Attention }-- Movement --{ Perception

SL: AMP is just the acronym for Association Model of Perception.
AMP also has in it the first letters of Attention, Movement and Perception.
Movement, response, action or inaction are the outcome of attention and perception.

I have heard neuroscientists say that the prime purpose for the brain is to control movement of the body. Think about it. What would you do and be, if you did not have a brain? If you could not move and program the brain with experiences of your environment? To instigate new movements? To learn from associations? A tree does not need a brain. It is genetically hard programmed to go for the sunlight.

Fitz: Can you give me an example of how my own mind works when it comes to the two modes of operation?

SL: I love to show you this page from Kahneman's book about 'Thinking, Fast and Slow'. In his terminology System 1 is the fast one and System 2 is slow.

Have fun:

The Lazy System 2

One of the main functions of System 2 is to monitor and control thoughts and actions "suggested" by System 1, allowing some to be expressed directly in behavior and suppressing or modifying others.

For an example, here is a simple puzzle. Do not try to solve it but listen to your intuition:

A bat and ball cost $1.10.
The bat costs one dollar more than the ball.
How much does the ball cost?

A number came to your mind. The number, of course, is 10: 10¢. The distinctive mark of this easy puzzle is that it evokes an answer that is intuitive, appealing, and wrong. Do the math, and you will see. If the ball costs 10¢, then the total cost will be $1.20 (10¢ for the ball and $1.10 for the bat), not $1.10. The correct answer is 5¢. It is safe to assume that the intuitive answer also came to the mind of those who ended up with the correct number—they somehow managed to resist the intuition.

Shane Frederick and I worked together on a theory of judgment based on two systems, and he used the bat-and-ball puzzle to study a central question: How closely does System 2 monitor the suggestions of System 1? His reasoning was that we know a significant fact about anyone who says that the ball costs 10¢: that person did not actively check whether the answer was correct, and her System 2 endorsed an intuitive answer that it could have rejected with a small investment of effort. Furthermore, we also know that the people who give the intuitive answer have missed an obvious social cue; they should have wondered why anyone would include in a questionnaire a puzzle with such an obvious answer. A failure to check is remarkable because the cost of checking is so low: a few seconds of mental work (the problem is moderately difficult), with slightly tensed muscles and dilated pupils, could avoid an embarrassing mistake. People who say 10¢ appear to be ardent followers of the law of least effort. People who avoid that answer appear to have more active minds.

Many thousands of university students have answered the bat-and-ball puzzle, and the results are shocking. More than 50% of students at Harvard, MIT, and Princeton gave the intuitive—incorrect—answer. At less selective universities, the rate of demonstrable failure to check was in excess of 80%. The bat-and-ball problem is our first encounter with an observation that will be a recurrent theme of this book: many people are overconfident, prone to place too much faith in their intuitions. They apparently find cognitive effort at least mildly unpleasant and avoid it as much as possible.

Fitz: Great! What was your answer?

SL: 10 cent

Note that you just had an experience of your inner workings, how your mind processed the bat and ball question. We perceive not only the world outside of the skin but also the world inside of the skin. We perceive our state of being, though it is tainted by the associations, which are dominant at the moment and which give us the context for the present state of being. Observe how you answer the question: "How are you?"

Note also that in general, perception works against a background, a context. Foreground versus background. A perception derives its meaning from the context. If it is pitch dark and you hear a noise that reminds you of a mouse, then you ask yourself: Does that sound make sense here in this fancy hotel room?

When I listen to the playback of a stereo recording and close my eyes, I ask myself: Where am I?
My enjoyment of the performance is greatly enhanced, when I do not hear left and right loudspeakers and the listening room, and only a phantom acoustic scene. If the perceived scene is 3-dimensional and naturally arranged spatially, and in a plausible environmental context, then I love to immerse myself most fully into experiencing the music. And that can get loud to a disinterested bystander.

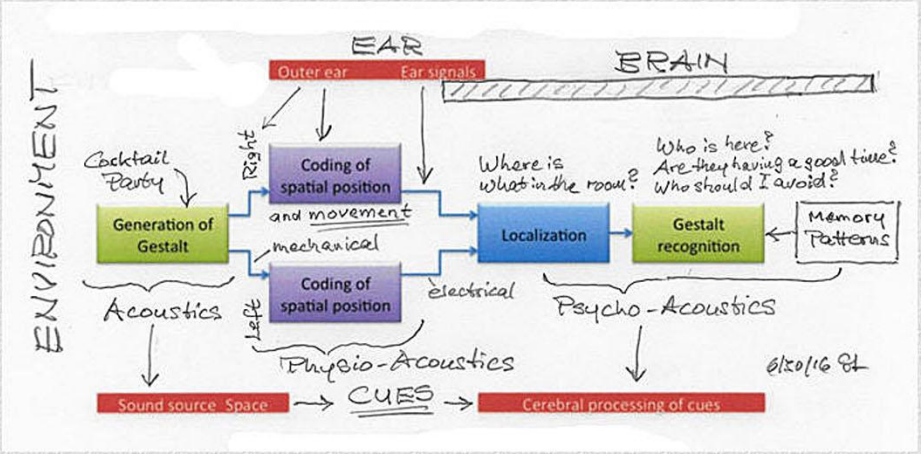
"Musik wird oft nicht schön empfunden, weil stets sie mit Geräusch verbunden", said the cartoonist Wilhelm Busch a long time ago.

SL: OK, let's move on.
I now I have laid the basis for talking about the specifics of hearing and how we process acoustic vibrations into meaningful perceptions. After my detours into general principles of perception, life and living we are back to how we are programmed to listen to loudspeakers.

SL: And before I forget I must point you to a book, which illustrates in numerous examples how the AMP is being gamed for financial gain everywhere you look: "**Phishing for Phools - The Economics of Manipulation and Deception**" by George A. Akerlof and Robert J. Shiller, 2015.

2.5 - Hearing - Part 2

SL: I have annotated the block diagram for the Association Model of Hearing to help explain it.
Let's assume we are at a cocktail party. There are groups of people and everybody is talking. That is a lot of noise or air vibration, all of which carries information. Since you always wanted to talk to Fred about your speaker project, you seek him out in the crowd of people, and move towards him, to engage him in a conversation. You call out his name, to get his attention.

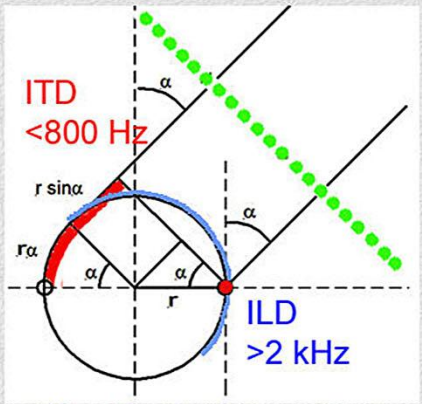


SL: By now being close to Fred you are in a different Gestalt of the cocktail party, than you were in, when you looked for him. But it is still terribly noisy and you have to turn your right ear towards Fred to understand him. Fred's voice and words have your attention and you positioned yourself optimally to discriminate against other voices.

By doing so your right ear receives a higher signal level and more high frequencies from Fred's voice than your left ear, because your head blocks sound from reaching the left ear. The electrical signals from your ears have been coded with spatial position of your head and with your spatial position in the room relative to all sources of noise/sound/voice.

By associating the spatially coded information from the ears with stored memories and learning, the air vibration Gestalt is recognized as Fred's voice and many people talking at a cocktail party. The brain also knows how to correct for the unequal signals at the ears from Fred's voice, so that he sounds the same in timbre, whether you look at Fred or turn your head.

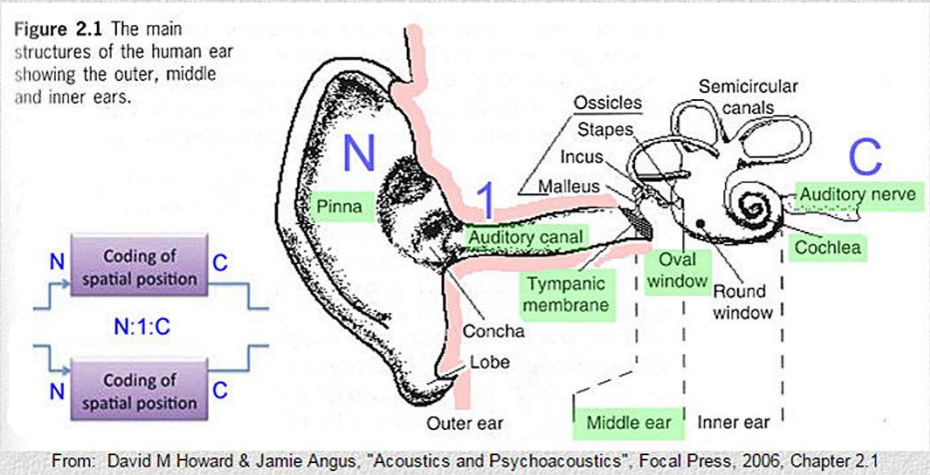
Such spectral correction of ear signals is important for tracking the location of a moving source, which has drawn your attention, relative to other sources of sound, which could be real or could be reflections coming to you from the environment of the moving source.



See: A model for rendering stereo signals in the ITD range of hearing

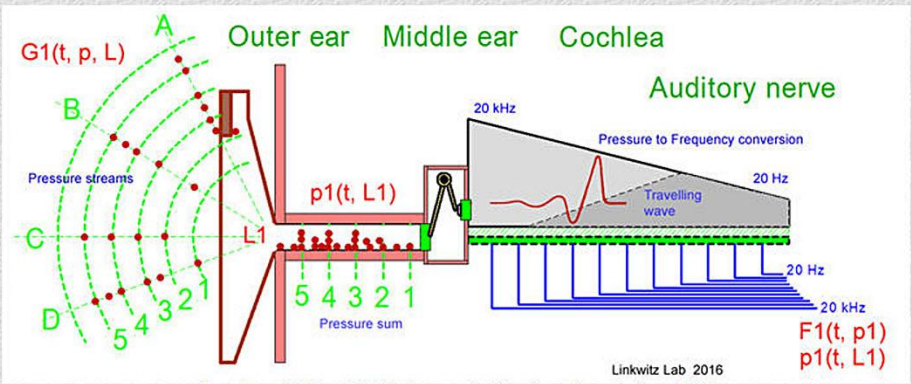
Turning of the head is used to determine the direction from which a sound is coming. The ears are spaced by about 17 cm = 2r, which corresponds to half a wavelength at 1000 Hz. That corresponds to 180 degrees of phase shift between ear signals. A signal, which reaches both ears from an angle a will have a phase shift between the ear signal outputs to the brain. Below 1 kHz the phase shift is unambiguous, but above 1 kHz a polarity reversal will bring back the same phase shift. Nature prefers the path of least resistance and so uses phase shift only below 1 kHz for finding direction. Actually, it is the inter-aural time delay, ITD, the group delay below 800 Hz, which is used. Above 2 kHz it is the ILD, the inter-aural level difference due to head blockage for higher frequencies, which is used to find sound direction.

The spatial coding that takes place in the outer and inner ear is quite fascinating.



From: David M Howard & Jamie Angus, "Acoustics and Psychoacoustics", Focal Press, 2006, Chapter 2.1

The pinna of the outer ear collects the N superimposed sound streams and combines them into a single stream in the auditory canal. The air pressure variations in the auditory canal move the eardrum, the tympanic membrane. The middle ear is essentially an impedance transformer, which converts the pressure variations in the air of the ear canal, to pressure variations in the fluid, which fills the Cochlea.



See also: J. Blauert (Ed.), Communication Acoustics, Springer 2005

Fitz: So far, so good.

SL: The drawing above shows four sound stream excerpts and coming from different directions, A, B, C and D versus time markers 1, 2, 3, 4, 5 and 6.
The dots represent pressure peaks of an acoustic Gestalt G1(t, p, L), which means G1 is a function of time t, of air pressure p, and of location L of the outer ear in space.
In the ear canal location L1 the four streams are converted to pressure variations p1(t, L1) by summation of the four streams.

The middle ear couples the ear canal to the cochlea by levers, i.e. by a mechanical impedance transformer, which increases the force generated by the eardrum motion to push with greater force on the fluid in the cochlea and to set up a traveling wave.

The wave travels over a basilar membrane, which responds to the spectral content of the wave. The beginning of the membrane generates electrical impulses in response to high frequency components and the far end of the membrane responds to low frequency content. The auditory nerve strand, which takes the electrical signals from the basilar membrane to the brain now carries both sound pressure information p1(t, L1) and frequency information F1(t, p1). The brain can turn the head and thus get another data set. From all that information and in connection with associations we recognize the Gestalt G1.

- **Gestalt recognition is based on CUES**
- **Two ears with fixed separation, and an obstruction between them, sample the sound-field of the original acoustic Gestalt.**
- **The sampling points can be turned and moved, thereby encoding the input data with spatial and spectral information cues**
- **The brain decodes the incoming data streams by association with cues from memory**
- **The recognized Gestalt is always SUBJECTIVE and is a sampling of the original acoustic Gestalt combined with ASSOCIATIONS**
- **What you hear is subjective**

Gestalt Generation and Gestalt Recognition

Hearing Process --- Algorithm, Strategy, Tactics --- Actions

1 Air pressure variations can carry information, which propagates out from a source in time and travels over distance, being progressively dispersed until the pressure wave is dissipated.

2 Information comes in a sequence of packages, which are streamed

3 Streams are collected by the outer ear and bundled in the ear canal. Ear shape, head and upper torso affect the collection

4	Behind the ear drum the bundled stream is converted into electrical impulses, which are then streamed to the brain
5	The bundled streams from the two ears are different, because the ears are in different locations in space
6	Turning the head changes the streams at each eardrum. The changes in eardrum streams relates the sound source location relative to where the nose points
7	The brain looks for patterns in left and right ear streams
8	Separating bundled streams into acoustic space background and foreground information Echoes are irrelevant for locating a source, but reflections must be qualified depending upon their direction and timing.
9	Having found a signature the brain assigns meaning to it by associating it with memory patterns and memory objects
10	If meaning indicates possible danger, then the brain calls for further attention leading to immediate action in the form of hiding, running or getting ready to fight
11	In order to act appropriately the brain must know the location of the danger, its direction, distance and movement
12	Comparing left and right ear streams for the danger's direction
13	Establishing/tracking direction and distance by head turning
14	Estimating size of sound object by its proximity and loudness
15	Recognizing the Gestalt by association with acoustic space, memory objects, feelings, emotions
16	Opening eyes: Confirming Gestalt by sight
17	Action choices: fight, flee, hide, surrender, relax, accommodate, enjoy

How do you respond?
Why?

GRACE
======= to respond without resistance =======
BEING

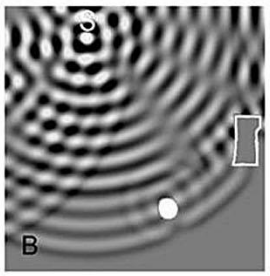
SL: Remember, earlier on we talked about the Zen master asking the student about the sound of a falling tree, about one hand clapping, and about the student asking:

"What is the essence of Buddhism?" And the master answered: "No Self, no problem!"

Fitz: Yes, and I said: "That gets heavy. Let's talk about hearing".

SL: Hearing is part of life and living. So if you ask me: "What is the essence of life and living?", the answer is: "to respond without resistance".

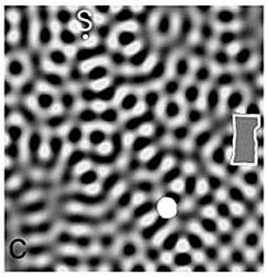
2.6 - In summary



Source

Generation of Gestalt starts

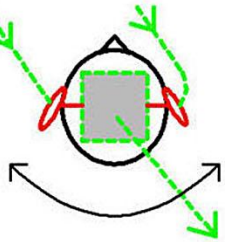
Receiver



Generation of Gestalt continues

Coding of spatial position and movement

Localization



Recognition of Gestalt continues in the Receiver

CUES and ASSOCIATIONS and MEMORY

Gestalt is recognized

RMAF 2015 - Linkwitz Lab

A cultural aspect to spatial hearing:

A cultural aspect to spatial hearing:

centered coordinates after compensation for eye fixation. Ultimately, subjects achieve a single external representation, with all sensory sources reconciled for consistency. The need for consistency implies some degree of plasticity.

Because the human brain is so adaptive, assertions about human nature or innate perceptual abilities are philosophically problematic. Such assertions fail to recognize the importance of culture: the microculture of the infant, the miniculture of the adolescent, and the macroculture of the adult. As a general conclusion, Dean V. Buonomano and Michael M. Merzenich (1998) explain: "the cortex can preferentially [re]allocate cortical areas to represent selected peripheral inputs. The increased cortical neuronal population and plasticity-induced changes is the coherent response... thought to be critical for certain forms of perceptual learning." Learning is an adaptive response to the environment; our brain is a manifestation of culture.

With regard to the impact of culture on auditory perception, Georg von Békésy (1960) reported an experiment in which a male Rom subject showed normal pitch discrimination but extremely poor loudness discrimination. Because his musical tradition considered pitch rather than loudness as being dominant, he could hear loudness but discarded it as having no significance. Thus not attending to an auditory attribute is, in effect, equivalent to not experiencing it, like irrelevant background noise. Just as a bushman, having lived in the forest for his entire life, would find it difficult to recognize and interpret the acoustics of enclosed spaces, so would an academic researcher, having lived and worked almost entirely in enclosed spaces, find it difficult to navigate the acoustics of a forest.

The problem in studying auditory spatial awareness is that the dominant aspect of learning does not take place under controlled conditions of a school or laboratory. Most learning is woven into life, be it listening to a mother's lullaby during the first nights of life in a nursery or attending weekly concerts and religious services in a church. In the school of life, it is usually not obvious what is being learned. By age 30, an individual has spent over 100,000 hours in a wide range of acoustic environments, which vary greatly across individuals and cultures. Olivier Deprès, Victor Candas, and André Dufour (2005) suggest that the improved auditory ability to localize found among those with myopia arises from the need to use auditory information during ordinary living. Rather than studying the biological properties of our species, scientists who explore auditory spatial awareness are actually observing culture. And depending on that culture, some individuals have more or less auditory spatial awareness than others.

From: Barry Blesser & Linda-Ruth Salter, "Spaces speak, are you listening?", 2007

Reading material and sound examples

Daniel J. Levitin, "This is Your Brain on Music", 2006
David M. Howard & Jamie Angus, "Acoustics and Psychoacoustics", 1996, 2006

Reading material and sound examples

Daniel J. Levitin, "This is Your Brain on Music", 2006
David M. Howard & Jamie Angus, "Acoustics and Psychoacoustics", 1996, 2006

Sound examples:
Ripping paper
Guitar
SFO streets
Symphony exit
Applause and inverted applause
Audio productions

Information processing

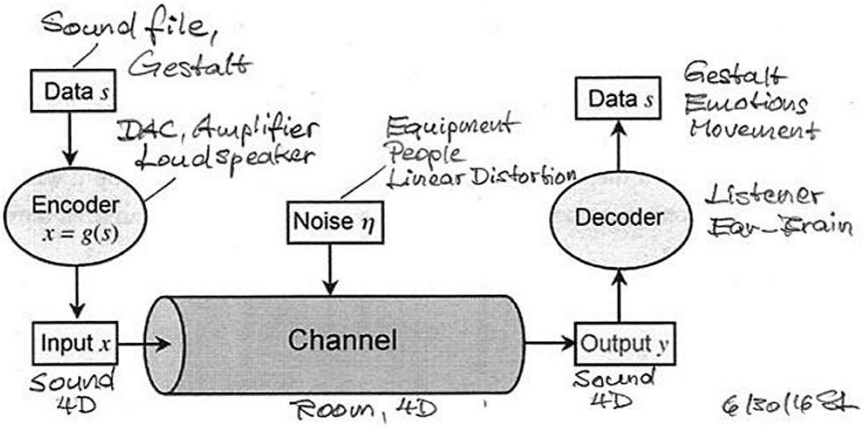
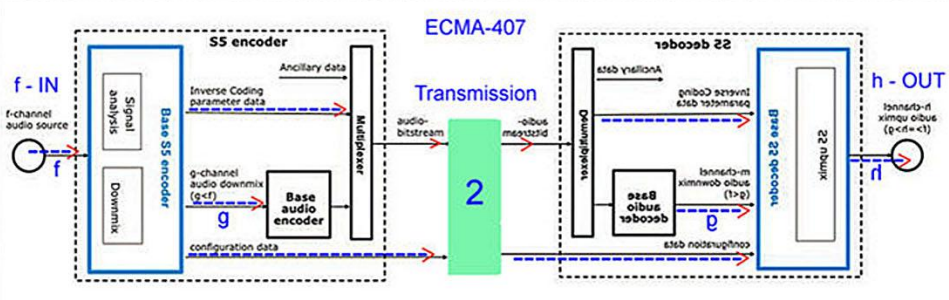


Figure 2.3. The communication channel. A message **s** is encoded as codewords **x** before being transmitted through a channel, which may corrupt the encoded message by adding noise η to produce outputs $\mathbf{y} = \mathbf{x} + \eta$. A receiver decodes the output **y** to recover inputs **x**, which are then interpreted as a message **s**.

From: James V Stone, "Information Theory - A tutorial introduction", 2015

From f-input channels via two bitstreams to h-output channels and with preservation of spatial information



From: Clemens Par, "Rationalism versus Empirism", 2014