GIANCARLO DEL SORDO

THE HEADPHONE WHISPERER

ACUSTICA

SIENNA

"Creativity is just connecting things. When you ask creative people how they did something, they feel a little guilty because they didn't really do it, they just saw something. It seemed obvious to them after a while. That's because they were able to connect experiences they've had and synthesize new things."

Steve Jobs

PREFACE

The idea of telling this story grew over time while I was working on software dedicated to the headphone emulation of physical space.

While this is a technically complex topic, my idea was not to describe the project in a scientific way. I think it will be much more useful to describe how the creative process can lead to technical solutions, and to explain how the birth of an idea can influence how software is created.

I think it's a topic that might spark the interest of someone out there, and a story that could serve as inspiration towards the realization of a project of any kind, not necessarily in the realm of audio.

I may not be the top professional in my field, but I've been lucky enough to have built tools that are used by many top professionals. I found myself bringing ideas to life in exactly the right place at the right time because I belong to a specific market. I'm proof that an average person with an exceptional desire to realize a dream can really make it happen.

THE JOURNEY

It was a February morning, and Antonio and I were almost at the end of the car ride to Viareggio. The cool morning air filtered through the slightly open windows, and it felt almost like a spring day.

We had left Lodi early that morning and were on our way to the studio of a dear friend, Gianni Bini, where several dozen headphone models from all over Italy were waiting for us.

Upon arrival Gianni welcomed us warmly, and after the traditional cup of coffee in the bar near his studio, he presented us with a box full of headphones.

"This is the first batch of headphones! Two more boxes should arrive sometime today, but tell me if this is what you're looking for."

"How many are there," I asked him.

"There will be at least fifteen, but I have more in the car and at the house. Oh, and I also have these Apple earbuds if you want."

In the box was an assortment of new and used headphones of various brands, all piled into a tangle of wires.

While we were discussing the plugin we were designing, Antonio was upstairs setting up the Gras calibration system on a small table in front of a couch, along with a portable converter and his notebook. The first set of headphones were already set up for measurement, and less than five minutes the initial data were already being processed on my laptop, which was resting on top of a console in a half-completed control room.

The first headphones we tested were an inexpensive model made by Presonus — I never would have expected them to sound any good. I prepared the calibration software, loaded the processed model, and then added our room emulation plugin, which was still an incomplete Beta version.

I couldn't believe what I was hearing. That cheap set of headphones suddenly sounded top of the line; in fact, it sounded like I was listening in Gianni's control room downstairs. If I closed my eyes, it felt like I was in a real recording studio.

The music was defined, with full low frequencies and detailed highs. It sounded incredibly exciting! But I wasn't listening in a professional recording studio — this sense of space was an illusion created within the headphones.

When the software was turned off, all the magic was suddenly lost. The headphones sounded like I'd have expected: cheap.

I yelled to Gianni: "Listen to our software!"

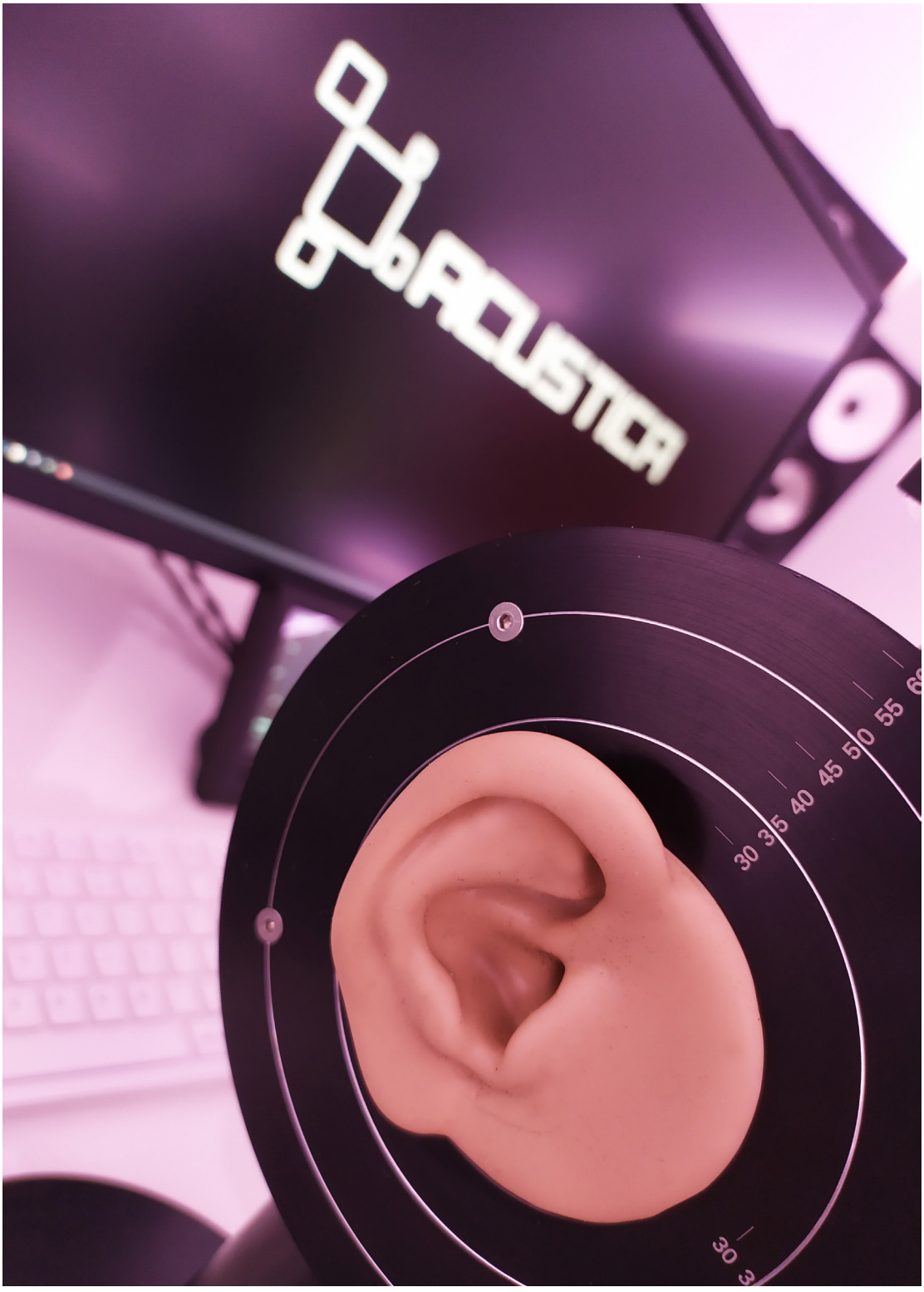
On the screen in front of him appeared an image of his recording studio, and from the headphones came the sound of a groove of a song that I knew well, as I had been listening to non-stop for several months.

Gianni seemed very impressed by what he was hearing, and two minutes later he had the phone in his hand. He was in love with the sound and wanted to those headphones in the studio at all costs.

The headphones were on loan —they had been sent to him by a producer friend from Rome. On the other end of the line, the distributor was telling him that he could have a pair in exchange for a small endorsement.

Gianni was right. In spite of the low price, those headphones sounded incredible.

The Sienna revolution was just beginning.


ACOUSTICA
30 35 40 45 50 55

UNDERESTIMATING THE PROBLEM: DUNNING KRUGER

As the CEO of Acustica Audio, I've always been involved in the digital emulation of analogue instrumentation.

For me, headphones have always been an analytical tool: a way to hear in detail the processing errors of our software, or to understand if the editing of a vocal track was correct. As an amateur producer of electronic music, headphones were the final quality control filter used in the last stage of music production, or a way to understand the specific quality of a synthesizer's sound.

Headphone drivers often have a very fast transient response; in other words, they react very quickly to short, percussive sounds, such as the clicks in a badly edited sequence. It's very difficult to hear this kind of subtle error while listening on speakers, unless you have access to very high-quality system in a properly treated room.

Speakers are usually "slow," in that the physical qualities of the room diffuse the effect of the perceived sound. The space you're in can create an auditory effect similar to applying a bit of reverb, which can dilute or mask any existing errors.

In spite of this, I've always considered headphones to be an unsuitable system for audio production, or at least not a definitive replacement for a good set of speakers.

Last year, we finally built a small recording studio control room in our offices. It cost a lot of money and was equipped with an extremely high-quality listening system. I was proud of it. After many years, I was finally able to listen to music under ideal conditions, which meant I could tell whether or not it had been recorded and mixed well.

One of the usual problems with amateur studios or listening rooms is that they don't have a flat frequency response or suffer from reflection problems caused by sound waves bouncing off of untreated materials. Obviously, you can get used to listening to music in this kind of environment, but if the room is not particularly "correct" in terms of its acoustic design it's difficult to get a good performance from your listening system, regardless of its quality.

For example, it makes little sense to have a sub — an amplifier dedicated to amplifying the low end — as a part of your listening set-up if the room reacts poorly to low frequencies. You'd risk making the listening experience more confused, and therefore useless.

Our new room sounded great, and in my mind, nothing could compete with it — not a different speaker system, and certainly not a set of headphones. It completely eclipsed the experience of listening at home, and I felt would never go back to the way I listened before. Finally, I could hear the reaction of our products at low frequencies and hear their defects and their strengths after having imagined the results for many years.

I'd checked them in recording studios owned by friends scattered throughout Italy and visualized their results with a spectrum analyser.

In the past, we'd encountered problems such as an artifact in the low frequencies that we had difficulty reproducing precisely because we lacked a listening system whose parameters would allow us to detect it accurately. Some of our customers could understand the flaws in our designs better than we could due to their superior listening systems. For this reason, we eventually decided to build a state-of-the-art control room, capable of accurately reproducing even the most subtle nuances, such as a complex bass sound or the expansion effect applied to the drums during mastering.

It was an October morning, and a few days earlier one of our competitors had announced their launch of a headphone system capable of imitating the sound of a high-quality physical listening space. The marketing was aggressive and promised that their product could perfectly mimic the sound of very expensive control room using a fairly inexpensive, high-performance headphone that he was introducing to the market.

I was looking for a new challenge, a new project to devote myself to, and the possibilities raised by this announcement grabbed my attention.

Was it really possible that the combination of a software system and inexpensive headphones could essentially replace a multi-million-dollar studio control room?

I ordered their product immediately so I could try it out at the office. I liked everything about the design. Steven Slate — the competition — had reproduced many different rooms, multiple listening environments, and the public discussion forums were on fire with the results.

To tell the truth, this was not the first product on the market of its kind. The company Waves had released a similar product, and a Russian product called Realphones had become quite famous in just a few months. Then there was the Sonarworks calibration system, which my friend Luca Pretolesi had been a raving about for years.

Another friend had often mentioned a product able to imitate the effect of listening through frontal speakers, called Canopener: the software solved the problem of excessive spatial listening in headphones, which is one thing that makes them an unequal substitute for a physical speaker.

I knew that this kind of audio leakage was called cross-feed. Because the speakers in a recording studio are positioned in front of the listener instead of laterally, the program had to imitate that same kind of phenomenon in a pair of headphones.

At the time, I was quite ignorant about all of thins.

Here's the thing: as much as I love audio, and even though I'm the CEO of company that works with audio, in many ways I'm not an audio expert.

As explained in the preface, the idea behind telling this story is not to provide a scientific analysis of a topic, or a technical solution to a specific problem.

It's about the creative process behind the project.

If I were to explain what led me to the creation of Sienna, I'd say it was primarily my enormous ignorance on the subject I was tackling; namely, of creating a sense of space while listening through a set of headphones.

I simply underestimated the complexity of the problem. I thought that it wouldn't be hard to do better than the competition, and that it wasn't necessary to solve all the problems in the audio world to create a useful product. Behind my ambition was a massive dose of unwarranted self-esteem. I simply wasn't aware of the difficulties that awaited me.

First of all, I had no idea that sampling a physical room would be so difficult. I didn't know that that something as basic to the process as measuring the parameter of a physical speaker was not a simple task. I had no idea that correcting the EQ curve of one headphone to make it similar to another was not a mechanical operation. And I certainly didn't realize that my competitors had already run into and solved other problems that weren't exactly trivial and had had plenty of time to work out the kinks in their designs.

According to the timetable I set, I had only a few months to make a product worth selling and I didn't know what I was doing at all. All I had was a little research I'd done a decade earlier on HRTF — the transfer function that accounts for how a human ear filters incoming sound — which had ended disastrously, prompting me to move on to other things.

After listening to the results published by universities such as IRCAM (Institut de Recherche et Coordination Acoustique/Musique), I decided to devote myself to something else instead.

On one hand, binaural simulation is a nice effect: by reproducing the result of the natural reverberation of a violin in a theater, as well as "placing" it in the three-dimensional space onstage, it creates an illusion that gives the listener the sensation of physically being in the audience.

On the other hand, the effect that is created is far from the audio quality necessary for music production or mastering, when the engineer needs to be able to hear minute details and therefore needs a calibrated, near-perfect system. It's one thing to recreate the acoustics of a violin in a theater, and quite another to convincingly simulate an expensive studio rig to the point where headphones could substitute for a top-tier professional recording setup.

The biggest problem is created by the unique shapes of our heads and ears. Every person's body possesses a different geometry: The shape of your pinna, or outer ear, is different from anyone else's . The ear canal, with shapes that vary from individual to individual, also vary in size.

These differences translate to an automatic equalization of the incoming audio, a sort of filter that our brain has learned to apply in order to compensate for these differences while translating the sound.

Thanks to the distance between our ears, we are also able to perceive distance and directionality. Sound is filtered according to direction, yet even when we listen through different listening sources, we have the impression of always listening to the same song. It's as if our brains were able to extract the directionality of the listening source and to break it down in relation to the source, allowing us to hear the original song as perfectly unfiltered.

But each of us hears filtered sound in a different way, and that way is determined by our physical form. Since it is very difficult to come up with a program that will compensate for the effect of each person's physical shape with precision, the risk is that each listener will perceive the sound in a "filtered" way. This is exactly what you don't want during the delicate phase of mixing and mastering a piece of music.

In fact, in that kind of professional setting the tolerances are minimal: an equalizer in the mastering phase is often calibrated for values that are dozens of times more subtle than the natural filter created by our ears, as well as the variable between our way of listening and that of our friend sitting in the chair next to us.

Our pinna filters sound for values around 10db, and probably in a different frequency range than the people around us. If we try to adjust an equalizer in the range of 0.2dB (a tiny difference in value), and the value of the filter is not correct, it will be impossible to evaluate the delicate intervention needed for the mastering of a song. The sound will therefore appear incredibly filtered, and therefore unacceptable.

This previous experience should have put me off the project — yet other developers were releasing products that users found interesting or practical, which meant they must have found a solution to these same issues.

As I read the praise for Slate's product, it occurred to me that the problem must not be so complex after all. He made solving it look like an elaborate but doable process; the users described his product as a workable solution.

But as I would soon discover, I was fooling myself into thinking I was dealing with an easily solvable task.

I think this is often the driving force that leads a person to experiment in an area that isn't exactly his or her comfort zone. The underestimation of the problem, and the challenge to dive into something you know little about, can be a potent combination.

Throughout my life, whenever I've achieved at least modest results when trying to solve a problem, it's because I didn't get discouraged. I believed the solution was close at hand, and so I kept at it.

This positive mindset can sometimes allow you to overcome enormous obstacles, because the illusion of your objective being right around the corner keeps you from feeling fatigued. It gives you the energy you need to spend hours analysing a problem.

This is what happened with Sienna. The rational side of my brain told me that the problem was difficult, and that I had already failed once. But the creative side kept telling me that if I succeeded, I would create an interesting tool for others, that the project would be fun, and that if others had been able to find some kind of solution, I would too. I had had everything I needed to succeed.

This kind of recklessness led me to dive in without having bothered to properly calculate the depth of the water, and to drag myself into an exhausting — yet rewarding — creative vortex. It would be a race against time to get the project finished on time.

However, simply copying the idea of a competitor is not something that appealed to me. The next morning, in our usual Zoom meeting with everyone at the office, I explained a broad outline of the project that I had in mind. The idea would be to copy the basic concept behind a competitor's product and improve it. For example, the product could be enriched with the reproduction of the harmonic distortion of the speakers, or with the minimization of artifacts, using some of the techniques we'd already developed.

The meeting ultimately took the form of a product manifesto.

I told them that we were the right company to embark on this venture because our specialty was sampling, and because we had the capacity successfully sample anything. Sampling is an activity that requires the coordination of a number of people: the people who reach out to the owner of the studio, followed by a team that physically measures the site, as well as a support team that follows these operations remotely and helps to solve the various problems that come up during the process.

It also requires the use of a framework that allows for fast prototyping, as well as tools that are as efficient and specifically calibrated as possible, such as scripts and custom applications.

After so many years of experience, our real specialty is to keep product costs down by shorten the time that is needed for prototyping and development.

At the time, we had just reached an agreement with an external consultant, Marco Vannucci, who was planning to devote himself to a project of sampling various physical spaces. Thanks to his advice, we had already purchased a set of microphones, preamplifiers, and converters that he would set up in various theaters and recording studios in Italy, traveling in a van set up for the purpose of measuring the data he collected.

It occurred to me to involve him in this new project as well, so I picked up the phone and called him.

"Listen, Marco, have you seen Steven Slate's product?" I asked.

"No, what is it?

"He's selling this new headphone design, and everybody who's tried them says that they're great at reproducing low frequencies," I told him. "But the interesting thing to me is the software: he's built a kind of recording studio simulation, meaning you get the illusion of being in a professional studio just by using his headphones. He says you can replicate the effect of acoustically treated very expensive rooms. The internet has gone crazy."

"That's interesting, but what would you like to do," Marco asked me.

"You know the project of sampling theaters and physical spaces? When we go to a recording studio, we'll have the studio director on hand. It's worth sampling those spaces as well."

"And how would you like to do that?"

"I don't know... can you come over next week? I already bought Slate's headphones; they should be here in a few days. We'll get together and work out a solution."

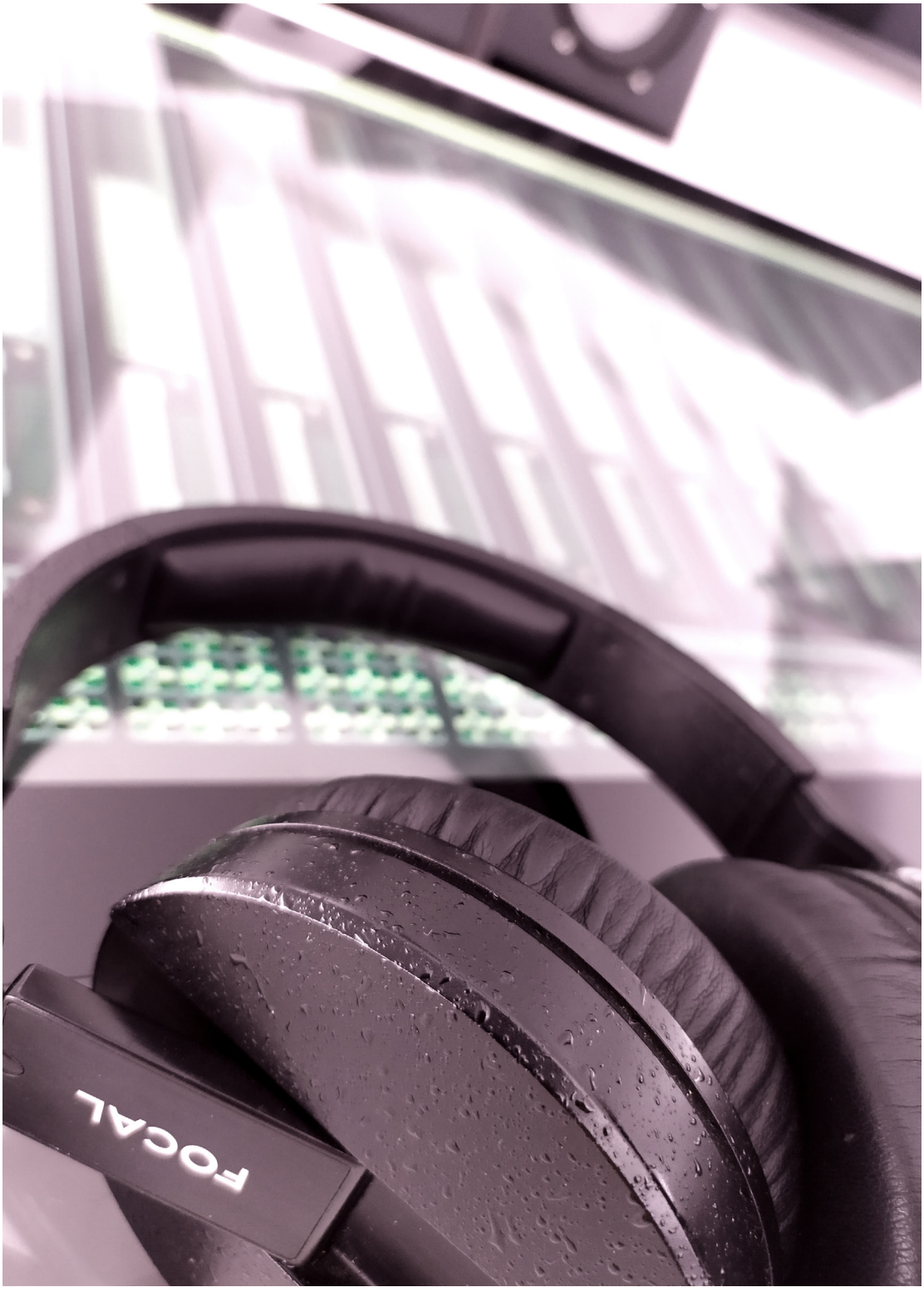
I wanted to start work on the project as soon as possible.

"Of course I can," he said, "but we need to figure out whether or not the final quality will be what we want. Usually, this kind of software doesn't work very well."

"That's what I've always thought," I said. "But others have done it. We don't even have to invent the best product out there — we just need to make one that creates that same effect, but in our own way. I'm sure we can figure it out."

That same evening, I spoke with my partner Umberto and convinced him to buy an expensive dummy head created by the audio experts at Neumann, which we would use to do our simulations.

I still didn't know how we would accomplish what I'd set out to do, it, but I was sure that the following week we would have come up with some at least a few ideas regarding a possible solution. My excitement was contagious — Umberto was also happy that we were embarking on a project a little different from the usual.


FOCAL

THE FIGHT OR FLIGHT RESPONSE

Steven's headphones were late and didn't arrive in time for the tests. We had a few audio files on hand and were struggling to get an idea of what the product might sound like. In the meantime, I had purchased software called Realphones, and was trying to figure out if the quality of Steven's product would be something similar.

Some users spoke very highly of Realphones as well, but when I tried it out using the headphones I had in the office, I ended up struggling to understand the music.

The sound appeared very filtered, almost like it had too much reverb. It was a great product, but it didn't produce the effect I'd expected to hear. Maybe the problem was me?

Slate's comments on the product in public discussion forums also contained points that I didn't understand.

Apart from the huge number of positive comments, there was a mention of out-of-phase sound for simulations mimicking listening in a car. That caught my attention.

Was there anything more I could do that they hadn't done? And more importantly, could it replace a very expensive listening system like ours?

I took a sheet of paper and explained my idea to Marco:

"Listen, imagine a person sitting in front of the speakers in a studio control room. If you listen through only the left speaker, the sound will reach the ears at different times. Maybe that's why people call the sound out of phase," I said.

"How about simulating a recording studio that has a frontal mono speaker?" I continued. "In the end, the mono signal is the most important part of a song, and it's what everyone wants to hear in the highest possible quality. Since our ears tend to naturally filter out this signal, it's probably not that important to filter it out in our programming. That way, we can achieve a higher quality sound than other have gotten so far."

We used this idea — partly right and partly wrong — as our starting point and began to create our first simulations.

One of the first samples was made with the help of Neumann's dummy head. But when we tried the result of our simulation on a song, it sounded terrible. Our disappointment was enormous. The sound was filtered, and incredibly out of phase — it was completely unusable. At this point, I understood that my evaluation of Realphones had been too hasty and that the product was almost miraculous compared to our first attempt.

Marco suggested that we try taking a measurement using the same technique he used to sample physical spaces. We had a set of microphones and a kind of tree stand that would allow us to position them precisely.

We did tests with various types of mics. The audio quality had improved, but it was still unsatisfying. As we experimented, we each worked up the courage to tell the other that the results weren't all that bad.

The problem was that it didn't sound like we were actually in a room. The sound seemed to be coming from an overly reverberant space, and mostly from the wrong direction. To make matters worse, the sound from the speakers sounded low quality, as if it wasn't linear. It was a chewed-up, unforgiving sound.

Marco had done his research better than I had and was talking about Harman and target curves — maybe that was the problem? I realized that I knew nearly nothing about the subject I was taking on, and that things were not as simple as they had first seemed.

After all, some people using Steven Slate's product were claiming that it sounded better than their own set-up. Meanwhile, we hadn't managed to achieve anything better than an awkward effect that only vaguely resembled the room that had been measured.

The body's fight or flight response doesn't only happen in the face of a dangerous attack that forces the subject to make a life-or-death choice. It happens every day, triggered by much simpler and comparatively low-stress situations, such as tackling a work project.

When you realize that you are about to embark on an undertaking that you don't know if you will be able complete, you have to decide whether to keep potentially wasting your time, or to give up. You have to choose whether to throw in the towel or face the music. You have to figure out whether the cost is worth the gain, whether it makes sense to continue to invest energy, or not.

We were faced with a question of whether we had the knowledge, preparation, and time to complete the project that we had set out to do.

A few years earlier I had developed an algorithm that equalized a signal to mimic the average curve of another reference song. This type of tool is called a "matching EQ" and had cost me many weeks of work.

Creating this algorithm answered an apparently simple mathematical problem, but one that could be approached in many different ways. At the time, it was developed for sampling tape recorders, which due to their slight compression tend to damage the slow signal used for the measurement, creating a filtered sound.

Later I would use this algorithm for the treatment of HRTF, though with poor results.

I had the idea to linearize the sound of the speakers in the room, as if I had the kind of hardware correction system commonly used in recording studios. The result was still very low quality, but at least it was a start.

We worked non-stop for three days, trying every possible combination of microphone placement and software correction.

Each evening Marco left the office exhausted, but we seemed to be making some progress. At the end of the third day, we were still not completely satisfied, but we had some ideas on how to improve the final result.

Marco has a much nicer control room than ours, with better acoustics and much more expensive speakers, so one idea was to try to sample the speakers in his room.

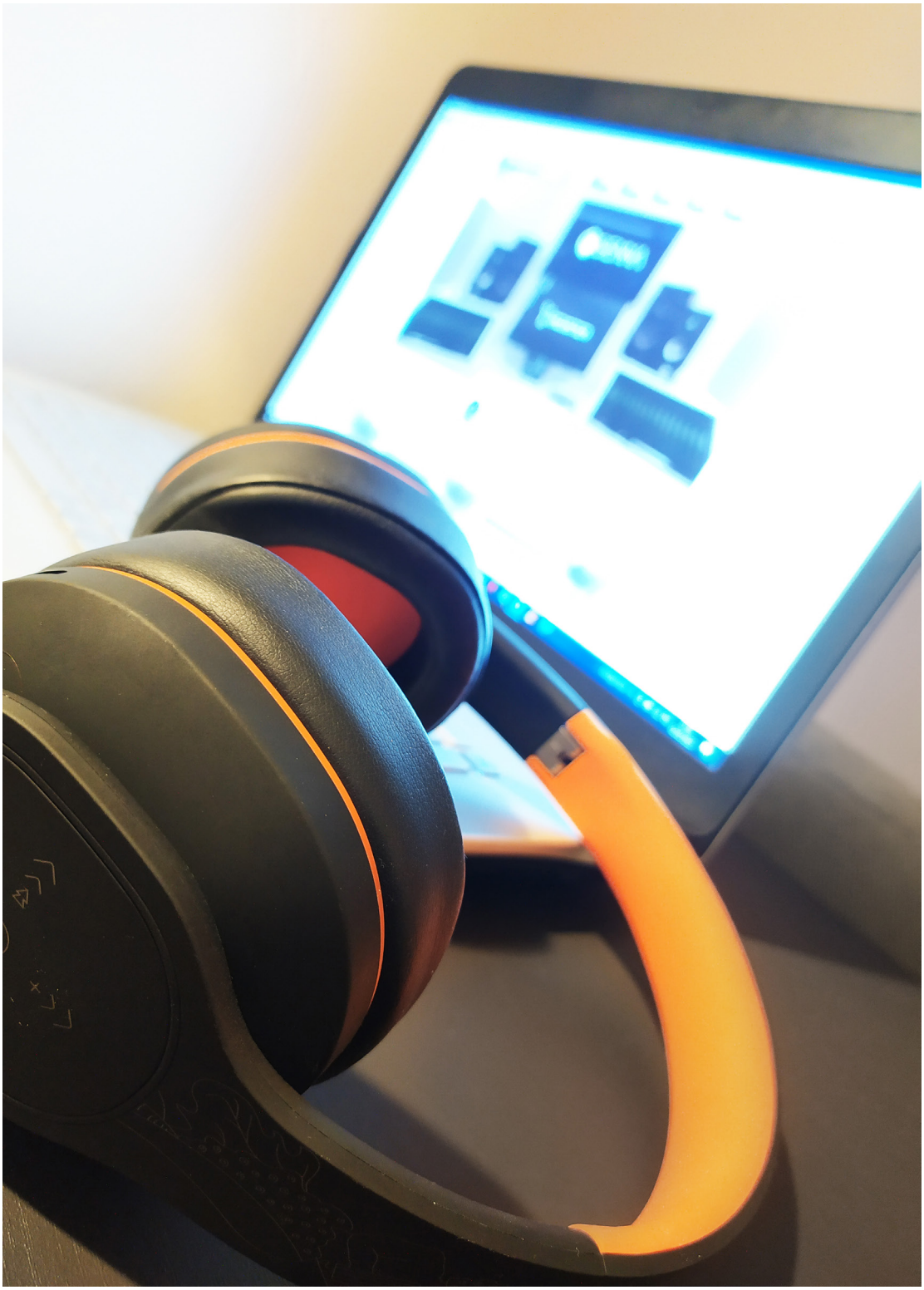
He said: "Your room has much bigger reflection problems than mine does. It's worth a try."

This could be a solution to increase the perceived quality of the middle component of the signal, i.e., the speaker that we were virtually "placing" front and center in the simulation. The side component or perceived difference, even if not completely convincing, could follow a different path. We could also make do with lesser audio quality, given its lower impact on the final result.

We left with the idea of giving it a try as soon as he reached his home in Pescara, which was several hundred kilometres away.

When Marco sent us the first results of the test, I immediately sensed a remarkable improvement. Maybe we had something viable after all, and after relatively few hours of trial and error.

It no longer made sense to run away — it was time to face the challenge and work to improve our knowledge. I spent the whole weekend studying everything I could find on the internet: books, videos, blogs. Maybe we really could do it. Now, the problem was to understand what it was that the others were doing and find out if they were really that much better at it than we were.



Finally, Slate's headphones arrived at the office. And no, they were not at all what I expected.

The headphones themselves weren't bad. Ok, the construction was a bit cheap, but compared to me my AKG headphones they had some pretty powerful bass. The low-end response was even better than the expensive model by Focal Clear that I usually used in the studio.

I didn't particularly like the high frequencies, and I would later learn that I prefer the sound of open headphones, even though their use is limited in noisy places. (Until that moment, I had never even thought about the difference between using open versus closed-back headphones.)

The headphones were fine. The software, unfortunately, puzzled me. My first impression was of an extremely filtered sound, and I immediately ran to the online discussion forums to search for comments that might confirm my impression. As it turned out, others had had the same experience.

If this was true, then why did some people seem to get a result that was good enough to replace their very expensive physical systems? From the photos they posted, it looked like they had costly, well-calibrated rooms, and these were people who usually paid a lot of attention to detail when it came to discussing plugins.

I phoned Marco.

"Marco, the headphones finally arrived."

"And... how are they?" he asked.

"I really don't understand," I said. "They suck, we won."

If I were to describe the reason why we decided to continue with the development of Sienna, it could be encapsulated in the ignorance we demonstrated in that moment. What we didn't know was that the product we were analysing was calibrated for people with a fairly wide ear canal, whereas mine is decidedly narrow.

Steven Slate would later release an update for people with a similar profile to mine. If he had done it sooner, I probably would have been discouraged, but circumstances left me with long period in which to enjoy what I felt was my moral victory. I was sure that I could create a better product, even if I didn't quite understand my own motivation to do so.

"I hear an extremely filtered sound," I told Marco. I really don't understand what's behind all this hype. It sounds like mixing on headphones when you have a cold. It's crazy, I could never work with an instrument like that."

The point is that if he could achieve a product that produced an undeniably good binaural effect — in spite of its also creating the feeling of having plugged ears — we could work on a similar program that included an algorithm that did not filter the sound and would therefore be useful in the early stages of music production.

In fact, if the translation of the mix (or the ability for the sound of the mix to be translated by various speakers) could be achieved with a product different from ours, we would almost certainly have solved the problem of sound design, as this is the phase in which the stylistic and timbral choices of a song are made.

But none of this was adding up. I started to measure the software, looking for confirmation of what I was hearing

I blew up Marco's phone with voice messages on WhatsApp.

"Look, I measured this stuff, and I still don't understand. There are peaks of about 10db, I've never seen anything like this. Why don't users complain?"

And Marco asked: "Yes, but how does it sound? Does it achieve the right effect?"

"The effect of being in a room is there, that part is well done," I admitted. "Maybe I've gone crazy, or maybe the software has gone crazy, I don't know. I can't understand why I'm the only one noticing this major flaw. If you read the reviews, hardly anyone mentions this problem, other than in a few isolated cases."

"Look, I'll bring the headset with me when we meet, and you can tell me what you think," I finally said.

In the office, I would spend my days testing the headphones with the rest of the team to get their opinion. Some would perceive a worse result, others a better one. I didn't know what to believe anymore.

A few days later I joined Marco in Ferrara, where we'd rented an anechoic chamber to allow us to precisely sample the speakers in our control room. The ideal environment in which to make the kinds of measurements we needed was a room without reflections, and we hoped to learn a lot from the results.

However, due to a problem beyond our control the room was only semi-anechoic: in other words, the floor was reflective. Luckily, this wasn't a difficult problem to fix with editing after the fact. We'd simply have to surgically eliminate that one reflection from the final measurement.

I brought Slate's headphones and couldn't wait to let Marco try them. I trusted his judgment; after all, he was the person with whom I had conducted much of my prior analysis. He had also followed me and supported me in my implementation choices on this project from the beginning.

Unfortunately for me, Marco has an ear canal that is shaped similar to mine, so Marco also heard a very filtered result.

Suddenly I had an idea: we could create a similar curve!

I phoned Stefano Dall'Ora: "Listen, Stefano, I'm sending you some software measurements. Can you create a curve like this one, even if it's not quite identical?"

Stefano is an incredible technician, and after less than twenty minutes I was already experimenting through the filter he had created on the spot. The idea seemed to work.

Unlike the reference software, we'd have made our curve parametric. A nice big dial would have allowed users to adjust its settings, so that if people needed to hear things a certain way, the control would allow them to choose the right amount.

But as a user, I would have avoided using the control. I wasn't convinced by the result.

TIME FIXES EVERYTHING

They say that time solves all problems, and so it was in my case. After spending the entire month of November analysing competitors' software, I was gradually getting used to the filtered sound. I had an initial graphic prototype of Sienna, which simulated Marco's room and which I listened to strictly with Slate's headphones, and with which I was making constant comparisons.

By now I was convinced that the filtered sound was a trick to get a better translation of the mix, as the exaggerated filter allowed the low frequencies to come through in a much more efficient way. In the end, the biggest problem in a mix is often the amount and definition of the low frequencies, but thanks to their exaggeration when compared to the mid frequencies in this context, it was possible to understand them better.

Intrigued by the progressive addiction of the reference curve, I began to contact possible candidates for a quick third-party check.

At the time, one of our beta testers named Simon Nakra was also a fan and a beta tester of Steven's, and he seemed to like both our products equally. My idea was to let him try Sienna to see if he would apply the filter curve or not, giving him the opportunity to adjust it to his liking.

I didn't ask for much — just his feedback on that one control element. To my surprise, he adjusted the filter curve to exactly to the amount present in the other software! Since Simon also has a very fine ear — a quality he often demonstrated in the field during the beta-testing of our products — his reaction was the spark that would finally help me start to see things clearly.

It was likely that the curve was similar to the HRTF of Simon's ear and very far from mine. But I could hear a convincing emulation of a room at a different setting, so I was happy to have a control that allowed me to calibrate the curve to my needs.

At this point, I began to research the origin of that curve, since it was advertised as an average value that was suitable for everyone.

I spent the month of November editing the curve according to my perception, putting together curves from various software programs and from academic papers, looking for a compromise that would give me a better-quality result. Every day I sent a new version to a control group, looking for feedback on how the curve was perceived, trying to discover the ideal compromise that would achieve the best overall effect.

I had also written software that allowed me to quickly derive filters from curves I found on various websites: the idea was to provide a curve that was as universal as possible, and independent of the physical characteristics of each person's ear.

We perceive the directionality of a sound based on a filter combined with the delay with which the sound reaches our two ears. Our physical conformation creates a filter, and our brain is used to processing this information. All I had to do was derive a reference curve that had a scientific basis but was sufficiently generic to function for everyone.

I often called Marco during these days, who confirmed that I was making progress in my search for perfection.

"I tried my Sennheiser HD650s and I hear a good replica of my room, it's an incredible result," he would tell me.

I knew I was on the right track.

The thing that pleased me the most was that I was gradually moving away from a copy of existing software and was approaching a personal solution to the problem. Having finally understood the mechanism behind it, I was free to create something of my own.

I believe that this process happens every day in every field: you start from a place of admiration or from the simple analysis of a reference product, but you begin to understand its basic mechanisms only when you start to pull it apart. After this process of deconstruction, you absorb every concept, you use it, and you end up sincerely appreciating the original to its core.

This is precisely when the importance of time enters the mix. Suddenly, this analysis combined with time allows you to make a creative leap, and to be moved by the momentum of your own decision making. At that point, we are in the hands of our instincts. We try to overcome our limitations, which can lead to encountering others which we then strive to overcome as well.

That's how I would sum up the creative process.


NEUMANN.BERLIN
STUDIO HEADPHONE NDH 20
45 50 55 60

THE UNPLANNED OBSTACLE

One thing we hadn't planned for was the profiling hundreds of different headphones. At first, I actually considered it, but then quickly discarded the idea.

After all, why would I embark on this additional, time-consuming undertaking?

There were already a few pieces of software that had repeatedly been described to me as extremely reliable and feature-rich; the best was probably Sonarworks, which had started the ball rolling a few years earlier by inventing a whole new genre of program.

Some of our friends had recommended the effective combination of Sonarworks and Canopener: the first one dedicated to headphone linearization, and the second dedicated to fixing the cross-feed problem.

Why does Sonarworks perform this linearization operation? In a word, it's a tool to make different headphones sound comparable.

It's logical that certain inexpensive models tend to sound different from models that are more costly due to the difference in material used in their construction.

On top of this, some manufacturers make the decision to try to reach a specific demographic by offering a timbre that is not scientifically superior but is designed to be more appealing to them. For example, younger listeners tend to particularly appreciate the brilliance of a converter and a transducer and tend to prefer it over a more linear system. It simply sounds more "pleasant" to their ears.

Studies have shown that some EQ curves, such as the Harman curve, tend to appeal to a wide sample of the general population. The reason is simple: the curve creates a result that recalls the way a speaker sounds in physical space.

In fact, the shape of the torso, the face, and the outer wear constitute a form of individual EQ curve that our brains have learned to translate.

When you wear headphones, some elements of our individual EQ are naturally taken out of the equation: one way to compensate for this is to artificially create a curve similar to the one you might experience in an average room in front of a pair of speakers.

We perceive the sound we've grown accustomed to as pleasant, so when something is missing from that sound, alarm bells go off. The sound could be described as good or interesting but would still be different from the experience of listening in a physical space. This is precisely the sound of listening through headphones.

My idea was to build an algorithm that could be used in combination with linearization software, and Sonarworks seemed like a good candidate. It was a solid product with a long history and enjoyed a strong adoption rate by the professional community.

I thought that if Steven Slate's headphones were considered to be sufficiently linear —in other words, that there was little difference when compared to the result after the introduction of his linearization algorithm — all I had to do was make my product work with them. That way, my product should automatically work with all headphones. I would just tell my users to use Sonarworks in combination with my product. (I already had a good excuse to do so, as most of my users had already purchased a license for the Sonarworks.)

As a sanity check, I tried Sonarworks with my AKG 612s to see if I could get roughly the same result as I heard in Slate's headphones. They weren't very far off: unfortunately, the AKGs lacked low end, but that made sense. Steven's marketing was all about low end, thanks to his patent-pending project and beryllium driver technology.

Over time, I completely forgot about the issue of different headphones. Marco was fine with his Sennheiser 650s, even without using Sonarworks. I was able to use my AKGs, and with Steven Slate's headphones my application sounded fantastic.

Eager to get a reaction from people whose opinion I trusted, I started to distribute the application to my friends, and collect feedback from them.

Among the first to try it were Paolo, the owner of one of the sampled rooms, and Alex, a talented producer of Italian music from Milan.

It was Alex's comment that struck me: "I tried Sienna, and I have to say that it works very well. However, there are some things that I don't understand; for example, how it makes my Ultrasone headphones sound. I've tried the Apple Airpods and you can hear everything really well there, but I can't make it work with the Ultrasones. Do you have any idea why?"

"But have you tried it with Sonarworks?"

I honestly don't remember the answer he gave me at the time, but he probably wasn't a regular user of that product. I suggested that he bring his headphones to me.

"I don't know, but I'm curious. Bring your Ultrasones to me, and the other headphones too," I said.

I had already guessed a possible reason for what Alex was hearing: if the Ultrasone headphones tried to mimic Harman's curve, they would have already amplified the 5 kHz band where our algorithm also had a boost due to the HRTF filter.

It was clear that some headphones needed particularly accurate correction, and that there were even some models that wouldn't work without it. The headphone profiling nightmare was just beginning.

So, I put my ignorance of the subject on display by proposing to Enrique that we measure the Focal Clears with the Neumann dummy head.

The problem seemed easy enough to solve: if the headphones exhibited non-linear behavior, it should be enough to capture their response and invert it in order to achieve a behavior closer to Slate's headphones.

But I wasn't ready for the next surprise. The measurement of the headphones indicated exaggerated boosts and cuts, contrary to what my ear was hearing — a rather linear headphone that didn't need any particular correction .

The confirmation also came from Sonarworks. The software presented a rather modest correction, extremely off-base when compared to what the Neumann measurement indicated.

It was clear that the dummy head was not the most reliable measurement tool to sample the frequency response.

THE UNEXPECTED SOLUTION

Over the next few days, it became clear that my strategy of inverting the measured curve of the headphones wasn't such a great idea. We had purchased a particularly inexpensive product, Minidsp, and were using it in place of the Neumann head. The measurements were more linear but trying to invert the curve still didn't seem to work particularly well.

I began to learn more about the subject, searching for solutions. In less than ten days I had developed an algorithm for where the headphones were failing, specifically in the area of 9-10kHz — where each ear has its own natural filter different from anyone else's — as well as the area above 10kHz, where the correction of the headphones is particularly problematic due to technical details related to their construction.

I was ready for Alex to take on the challenge using his headphones.

He came to see us one morning in December, curious to try the speakers we'd installed in our new control room.

"I brought you the Ultrasones. Try them out, everything sounds extremely bright," he said.

I tried them immediately. He was right, the high frequencies were totally fried. The sound was completely out of control.

Antonio and I immediately took a measurement and tried to apply our correction algorithms. But the result was pitiful, and it was clear that I still had a lot of work to do.

In the previous few days, I had found a way to download a huge database of headphone profiles from the internet and convert them into a format that could be combined with our measurements, to create a kind of average using our values.

The results were all over the place, thanks in large part to the different measurement methodologies that had been used to create the profiles. Some had used expensive professional systems, such as Gras and Audio Precision, while others had made them using homemade dummy heads. In some cases, an arbitrary reference target curve, like Harman was added during the measurement, and in other cases no curve was applied at all.

What was I doing wrong? Some headphones seemed to react extraordinarily well, others much less so.

I began to think hard about the system used for measurement. The Neumann head seemed very reliable when measuring low frequencies but became completely unusable in the high frequencies. In fact, the ear canal in the dummy head was completely absent.

In Minidsp, where the ear canal was barely noticeable, the measurement appeared to be much better, but that still didn't seem to be enough. Some headphones — for example, models that are designed on the earcup — seemed to exhibit fewer problems, while the ones that surrounded the earcup seemed to be more problematic.

In the end, I learned to listen to a reference track and compare the results to what I was getting out of Slate's headphones. Hearing the same song repeatedly while using different models gradually allowed me to become faster and more accurate in evaluating each individual headphone's results. Everything came down to memorizing the timbral content of a single song.

Step by step, my algorithm began to outperform Sonarworks' algorithm. This was not necessarily a great result, considering that I was profiling the headphones that were right in front of me, whereas Sonarworks had to be able to correct a generic headphone model that might be very different than mine.

Still, the improvement I was getting day after day was encouraging. Perhaps it would be worth measuring and providing a correction for "all" headphones after all, not just for a few. Of course, that would imply an immense amount of work.

Would it be worth it? The answer came in December, out of the blue.

One of our beta testers, Francesco Campbell, contacted me to let me hear a pair of headphones that he thought were particularly interesting, the Audeze LCD-X.

"I want you to hear these headphones, I think they work very well with your software," said Francesco.

Meeting with him appealed to me for two reasons: first, he was a Canopener expert, and I was interested in its comparison with the cross-feed implementation in Sienna. Second, I was interested in a direct comparison with another rather expensive headphone model, the Focal Clears.

He came to visit a few days later, carrying a couple of bottles of wine and a pair of headphones that at first glance looked gigantic.

"No wonder they sound like a set of speakers, they're as big as speakers!" some users were joking on the internet.

The first time I listened to music through them I was amazed by the translation of the low frequencies. It was like having a real subwoofer in front of me.

"You didn't tell me they sounded so good with Sienna!" I told him.

The result was already convincing using Sonarworks, so I felt it was worth trying to profile them with our prototype. I followed the same procedure as when we measured the other headphones, asking Antonio to get a set of measurements from the internet, preferably based on a better measurement system than ours.

My intuition told me to ask him to limit the search to one system — for example, Gras — and to combine it with our measurements.

The result was amazing. At that moment I understood that the use of a less-than-optimal system like Minidsp had led me to improve the algorithm beyond expectations. At that point it would have been enough to use a better measurement system, and I would have been more than happy with the results.

If I had had a well-functioning system from the start, I probably never would have worked so hard. My complacency would have led me to give up much earlier, and to consider the result I that had already achieved to be sufficient. The end result would have been good, but not great.

That day we did three things.

First, we purchased a set of Audeze. After hearing how well they worked with Sienna, we thought it would be a good idea to have them in the office to show to guests.

Second, we began inquiring about purchasing Gras. It was the ideal system to use with our correction algorithm.

Third, I asked Francesco to involve a friend of his, Cesare Marocco, who owned Audeze LCD-Xs. I wanted to see if he would be interested in using our correction algorithm. I also needed to figure out whether or not the result would be the same on his personal headphones, which were certainly different from ours.

That same evening Cesare, whom I didn't yet know personally, called me. He was enthusiastic, even excited. According to him, our profile sounded amazing.

The next day, I explained the new situation to the team: "I've decided to measure all the headphones with Gras. This means a lot of work, as I want to conduct this activity personally. But our algorithm looks really promising!"

"We have three months, maybe less, and I'd like to profile as many models as possible," I said.

Everyone started phoning distributors, stores, and personal contacts, trying to get their hands on every model on the market.

VALIDATION IN SCIENCE

An external observer might consider the creation of a new audio plugin as based exclusively on scientific foundations.

This is generally true, and very often the final product design is a result of exactly that. But the audience using the plugin does not need a scientific tool, they need a creative tool that allows them to complete a job.

For example, take the configuration of speakers in the control room of a recording studio: if we were talking purely about science, the EQ curve would be reduced to a flat line, which is what room and speaker designers in general try to achieve.

I say "in general" because not all sound engineers want a perfect, sterile, operating-room-type environment in which to mix or record. Many prefer a configuration similar to what an end listener might have, who will probably listen through cheap speakers placed in rooms that are certainly not designed for ideal sound reproduction.

When we started sampling the speakers, we were faced with an important question: would it be better to reproduce the speaker in a "normal" room with all its typical imperfections, or to strive for a final response that is closer to an ideal listening environment? The answer was anything but obvious.

I thought long and hard about it, taking into account Sienna's target audience.

It was clear to us that some users would prefer the sound of a room that mimicked a typical real-world configuration, with all its merits and flaws. If a mix sounds good in that physical place, it means that under professional conditions it would sound even better. However, some might be accustomed to a very different configuration, and would therefore be comfortable with a more neutral result.

One effect I have noticed during headphone listening is the continuous comparison the user makes between the original track and the one processed with the addition of a room simulation. In a physical space we can move our heads continuously, we can walk around freely, and the variability of the resulting sound is remarkable. When listening through headphones, this variability does not exist.

We've developed a habit of being more forgiving of listening defects; for example, the fact that the sound coming out of the speakers is not crystal clear in the high frequencies, or that the sound is incomplete or unbalanced in the low frequencies. It may sound absurd, but it's true.

Our ears continuously receive a filtered version of the sound, and our brain modifies this perception by automatically ridding it of the equalization curve that provides us with the spatial position of the sound source. In a nutshell, we believe that we are always listening to the same thing in the same way; but we are not, because we're not standing still.

This kind of experience could also be achieved with headphones, provided that we have a perfect version of our personal HRTF, which could possibly be obtained with a cast of our earcups.

But we have learned over time that this perfection does not exist. Everything, even the geometry of the individual ear canal, can affect the final result., and trying to support a calibrated HRTF tailored to a specific user is impractical for this kind of instrument.

In the end, Sienna's goal is to provide a tool that helps in the translation of a mix. Its purpose is not to scientifically recreate a personal listening situation, but to help engineers to complete a job. If you need a scientific tool perfectly attuned on the physical geometry of each unique listener, the customization costs would be so high as to make diffusion of such a tool impractical.

We've thought about solving this problem by exploiting cutting-edge techniques, such as the use of deep learning (i.e., neural networks). You could take a picture of the user's ear and use that image to derive a custom filter.

Unfortunately, this type of analysis will always be incomplete due to its very nature. Even if you could reconstruct a perfect version of the geometry of one user's ear, variables that would further complicate the analysis would be excluded from the calculations, such as the distance between the ears or the shape of the torso. All of these things affect the sound.

And if one wanted to take a truly scientific approach, he or would have to take into account the first layer of sensory neurons (the sensory cortex), which is the part of the brain that processes sound in combination with the eardrum. In short, we would have to measure the "downstream" system, taking into account the many pathologies that contribute to our sound perception.

I believe that we will soon be capable of creating a process that does exactly what I've described. The integration of biological neurons and artificial neurons is an increasingly popular field of research, so I don't rule out the possibility of finding a perfect, personalized, affordable solution in the not-too-distant future.

Until then, every possible solution is nothing more than an approximation of the final result. For not, we should ask ourselves a question: what are we trying to achieve? What problem are we trying to solve?

The user often has a very trivial problem: he or she works in a poorly treated room with imperfect speakers and would like a solution that helps to get a better sonic result out of the headphones. That's all.

Taken in that context, every improvement is an achievement. From this point of view, offering the experience of an "ideal" room, or one that is corrected of even some of its useless defects, is already quite helpful.

From the earliest stages of the project, we tried to correct speaker and room responses in order to offer a linear situation, a bit like a room correction system. But some hardware manufacturers like Trinnov Audio provide also provide practical solutions.

When we began planning the measurement of the control rooms, I suggested to Marco that we also consider using one of these systems in the field. We decided to buy one and test it on one of the first samplings we had on the agenda. The ideal candidate was the Fonoprint mastering studio — an iconic facility where a number of famous Italian songs were born — which was equipped with an extremely expensive listening system.

I reached Marco in Bologna one morning at the end of November. Fonoprint's studio is located in the historic part of the city and is equipped with security cameras that monitor the vehicles that come and go. Visitors can enter the premises as long as they park in one of the paid garages that validates parking.

Unfortunately, the parking garage I had planned to use was closed due to construction— constant construction is pretty much the order of the day in Italy's historic cities — so I was forced to wait until evening, when we could access another recording studio, Impatto Studio, which had a similar garage next door.

I decided to clear my mind and savour the moment. What was supposed to be a brief check of Trinnov's field operations had turned into a moment of reflection, which might indirectly lead to the solving of another problem.

Studio owner Marco Borsatti welcomed us to Impatto with open arms, in spite of our impromptu visit. While members of the team started measuring the room, he and I started to discuss the binaural issue.

Marco had a Neumann dummy head just like we did and was fascinated by the subject. We ended up talking about linearizing his room, which included Yamaha NS10 speakers with a subwoofer. He explained that he couldn't move the room to a more convenient location in the city because of the unique physical and sonic characteristics of his space, to which he had grown accustomed.

Would it have made sense to provide a linearized version of that specific situation? Probably not, as most users would have appreciated the original sound. Measuring with Trinnov wouldn't have worked, so I had to find a different solution. I decided that it would be much better to provide a tool that allowed the user to choose the amount of linearization.

I started to work on a linearization algorithm that would allow the user to customize the amount of the effect that was applied.

This didn't work well at all with the measurements obtained using Trinnov due to phase corrections made by the system: its final measurement mixed with the results created without it, indirectly creating huge interpolation errors.

Creating a custom linearization system would have allowed the problem to be solved in an elegant way, optimizing this crossfade feature. In order for this to work, it would have to be implemented for the speaker measurement as well as for the room measurement. Integrating the two measurements, making the linearization a parametric function, and forcing all these elements to coexist would have been an impossible task otherwise.

I worked on the linearization algorithm until Marco Vannucci joined me again in Lodi in mid-December, ready to take a new measurement of our room.

In order to verify the validity of the improved correction, I subjected the result to the most demanding standards. The process had to seem transparent, almost invisible even when using the most extreme setting, and reproduce a sound similar to that of the room's speakers. Everything had to appear fluid, natural, and minimize any artifacts as much as possible.

Marco tried the sound of the Impatto studio in his headphones at the extreme setting.

"Listen how good it sounds," he exclaimed. "This is the original room!"

We listened to the new version through headphones. It was less than perfect but extremely musical. And best of all, you could choose the amount of intervention to apply!

I had indirectly solved the initial problem. Although the direct comparison of the headphones versus the room would have been disorienting, by using the intermediate setting it was possible to maintain some fundamental characteristics of the room without producing a low-quality sound in the headphones.

We were very happy.


LEFT
UNCONNECT
CHASSIS TO EARTH

Though we were making progress, the development of Sienna's binaural features had its ups and downs. It seemed to be a much more complex problem than I'd expected and was also almost undocumented in regular literature.

In the first days of October, I asked Stefano to create a prototype in a commercial sequencer, Reaper, in order to experiment with the deconstruction of the room response into a component dedicated to the monophonic signal (the virtual central speaker in front of the listener) and a component dedicated to the side signal (the subtraction between the original stereo signal and the mono component).

My theory had a logical origin: monophonic sound has no directionality of interest and is usually played by a frontal speaker in the most expensive recording studio setups. I wanted to include a mono component of the highest possible quality because it's usually the most consistent part of the signal.

Everything else, often present in very small amounts, could be processed even with a less than perfect room response. The MP3 format usually has a rather poor encoding for the "side" part of the signal, but we don't usually notice when listening to a song in such a compressed format.

The strategy seemed to be optimal to solve the problem of the out-of-phase sound. If a mono signal also comes from the side speakers, the monophonic compatibility of the resulting signal will be compromised, so the simplest solution was to have it come only from a central source. That way, we could produce a signal that reached both the user's ears at the same time.

I suggested to Stefano that he process the monaural signal with only the speaker's response and leave the room response out of it.

The idea looked great on paper but ended up being more complicated in its prototype implementation. Stefano had spent an entire weekend translating my theory into practice and had corrected it by ear, finding a sort of compromise between what I claimed would work based on my intuition and what he perceived as plausible according to his sensory experience.

The following Monday I tried the prototype and it seemed to work pretty well. My yardstick was our previous attempts, and this most recent version seemed like a huge step forward. Sure, the Slate or Waves algorithm sounded more spatially immersive, but we had an extremely clean and solid sound on our side.

I asked Antonio to create the first version of Sienna based on that prototype, but the process turned out to be a hell of a conversion. Stefano's implementation was arbitrary, and every time I tried to modify it by moving it in one direction or another, I ran into new problems.

Stefano had processed some of the monaural components of the signal with the "side" room component of the sound, which meant that each time I decreased the amount of room sound I got a more pleasing result but with a less certain spatial origin. Every time I increased its amount, I got a sound that was too reverberant, but with good directionality.

On top of all this, Stefano's arbitrary composition did not seem to make any numerical sense, which worried me. Did it actually solve the problem of translation in the majority of cases? In spite of apparently lacking a scientific basis, would it be accepted by our users?

I tried to get a sample of our reference users to test the latest version of the algorithm.

Luca Pretolesi, my partner in a Las Vegas software company, still found the sound to be very out of phase. Ironically, the problem I was trying to solve was also the most obvious one. He has a very fine ear, and he was having difficulty working with this new implementation.

Out of politeness, he eventually stopped talking about it. I knew when he did that it meant he wasn't fully convinced.

I decided to try completely eliminating the mono component introduced by Stefano, and suddenly the signal was airy, clean and all the problems seemed to be solved. Mathematically, it all made sense, as that signal didn't have to exist at all.

But as it would turn out, I would have to prepare myself for another cold shower.

Robert, a tester to whom I had sent the prototype, found the translation of the stereo image strange. The same thing happened with another tester, Matt — his response was that that the stereo plan was too narrow.

But the coup de grace came from Alex. He called me at the beginning of December.

"I tried Sienna in a real work situation, and I must say that it's doing pretty well," he said. "However, I've come across a rather strange problem."

"Tell me," I replied.

"Today I was trying to balance the sound of an acoustic drum kit and I didn't quite understand how to use the pan."

The pan is the control that allows you to decide the proportion between the left and right signals of a channel.

In short, Alex couldn't place the elements spatially, which should have been Sienna's strong point. A physical speaker system has exactly that advantage over headphone listening: a spatial element.

I reassured him and experimented with adding more parameters to the algorithm we were using.

It seemed the problem with this implementation was how to translate signal strength information. A signal present in only one of the channels had to be converted into time information, so that one of the two signals reached one ear before it reached the other. Technically, everything was correct: if we exclude the equalization of the HRTF, realistically the signal of the left speaker reaches both eardrums at more or less the same time.

When dealing with a physical speaker system, it is the HRTF filter that created the magic, because the signal reaching either eardrum is much more filtered. This results in a different sound pressure intensity for each ear.

Conventional stereo recording techniques based on panning translate this effect into a practical approximation. The pan does not actually filter the signal, but instead works on the volume balance. This results in a high-quality, plausible, perceived effect.

The simplest solution seemed to be the partial restoration of this volume imbalance.

We debated the results we had in front of us versus what we were trying to achieve. Alex had partially solved the problem he'd run into during his session, but the signal still didn't seem to be coming from the right direction. I spoke to Marco and asked him to calculate the time delay between channels as precisely as possible.

As it would turn out, his system did not have microphones placed exactly 18 centimetres apart, which is the average distance between the two ears on a human head. The distance, due to practical considerations when it came to positioning the microphones, was slightly greater.

Marco set about applying his high school trigonometry, did his homework, and emerged from his calculations a few hours later with an accurate timetable charting signal delay.

Things were getting better, but we were still far from perfect.

I proposed making the delay parametric so that the user could virtually position the speakers as needed. While the idea eventually worked, at the time it seemed a bit improvised and without a real scientific basis.

Then one Friday at the end of January, everything changed. We had complicated the algorithm so much that we didn't even know what we had put into it. The mathematical process looked like a tangle of arbitrary equations.

"Antonio, we have to print out all the functions and try to understand what we're doing," I said. "I don't understand anything anymore."

I was reminded of a friend whose studio partner was completing a mix on a song in the late 90s. They had a digital mixer, and some of the processes were disguised by an unintuitive interface. In analysing the final configuration, he had discovered that the same channel would often have one action assigned to it in one direction and one in the opposite direction, and the two actions would end up cancelling each other out.

As it turned out, our algorithm was not much different.

"We have everything we need in here, and it sounds pretty good. We're probably just complicating things for no reason," I added.

This kind of approach was familiar to me, as I had adopted it countless times in the past. I started with the traditional theory, and by following my instincts arrived at a solution that seemed to sound better according to my personal parameters. At that point, I studied what I had in front of me and try to understand why it better answered the question at hand. Then, I proceeded to simplify the solution until I arrived at a scientific and replicable result.

So, what were we really looking at? We spent the afternoon analysing all the functions we'd printed out, studying them minutely.

As it turned out, our algorithm was nothing more than an attempt to sum the signal of the mix processed by the speakers with the binaural reproduction of the room. We had spent so many months on the problem, and all we had managed to achieve was to work around it!

But this realization opened the door to a particularly elegant solution: we would create a control that would smoothly transition from mix playback through the speakers to binaural playback, while also offering all the nuances in between. This type of control was particularly powerful: in some positions, the sound was qualitatively superior but more congested in the stereo plane. In others, it was airier and provided a better binaural illusion, although with a less precise and direct effect.

If we had started from a place of a purely scientific development, it's unlikely we would have ever come across this kind of solution. That's when the words of Steve Jobs came to mind: "When you ask creative people how they did something, they feel a little guilty because they didn't really do it, they just saw something".

Creativity is sometimes little more than an accident. The creative process is often simply the pull exerted by a crazy intuition.

We are not creative by choice, but by mistake.


Bowers&Wilkins

THE HARD WORK

What awaited me turned out to be a titanic job! It definitely merited the tongue-in-cheek term "headphone whisperer," which started out as a joke in our office.

The Gras system arrived in early January and marked the beginning of a small revolution. It worked well right out of the box, completely exceeding our expectations, but it also opened the door to new questions. How many sets of headphones would we be able to profile before the product was released?

We drew up a plan, taking into account all of the important elements of the market. A few days earlier I had even done a little survey, trying to come up with a list of must-have models. I never thought that this process would be such a crazy undertaking.

The headphone models typically used for music production is approximately a few hundred in number. Out of all of these, the most relevant can be reduced to just a few dozen, and the market penetration of many models is linked to their geographical area. For example, if one were to enter any Italian recording studio, he or she would probably find the exact same brands, but this list would not be identical in Germany, or in France, or any other country.

How could we address that problem?

Initially, we contacted our Italian friends and asked to have the most popular models sent to us, such as those made by Sennheiser. Marco had brought me his HD 650s, and I had used them to calibrate the correction algorithm.

In the office, we already had some Focal, some Sony, and several types of earbuds, and during this first period were able to sample the most popular headphone models in Italy.

I particularly loved the Audeze, but the Sony headphones weren't bad either. And some Bluetooth headphones behaved well enough, such as the Bose.

In the first phase, we identified what would work well with our algorithm and what would not. It turned out that there were two issues that needed to be addressed: the infrasonic range and the range related to medium-high frequencies.

Headphones often need corrective action in the low frequencies, but this is difficult to achieve from a practical standpoint. I began to realize that I needed to listen to each model, one by one, to see if I could reduce the intervention towards a more sensible and usable correction. Sienna needed to have a specific reference curve, but some headphones were not able to reproduce it correctly; for example, Beyerdynamic or AKG, as they often have strong limitations in this frequency range.

This doesn't mean that these headphones are of lesser quality than the others; it just means that they were less suitable for our final purpose, which was to reproduce a system of physical speakers, complete with a subwoofer. For example, the Beyerdynamic 1770 is a very good headphone: it has sold construction, is comfortable to wear, and offers a very detailed listening experience, regardless of the correction applied.

Unfortunately, this headphone shows a poor response in the low frequencies, and when you try to fix it, it actually sounds worse. The solution? You have to limit the amount of intervention.

The exclusion of a 400-euro headphone from the Sienna algorithm would not make much sense. It would be much more logical to create a more limited correction, which would allow for a good compromise between what is missing and what the system can deliver.

Another rather critical frequency range is at 5 kHz. This is the area where the generic HRTF provides a small boost, but it's also where some headphone manufacturers have implemented an artificial boost in order to make the sound more appealing

Since some individual examples of certain models can vary widely due to imperfections in the manufacturing process, we can't know if all the headphones produced by a particular brand will have exactly the same response or not. This is a problem, as it prevents us from using very steep or selective filters in the correction algorithm. A different factory could create a product with a slightly different response, and an excessively steep filter could cause more harm than good.

Therefore, the program has to use very general curves. When a headphone misbehaves in a particular region — for example, if the resulting sound is excessively harsh — the best solution is to exclude this region gently by using only slight attenuation. If this is not done in critical frequency ranges, it can result in disaster. Sienna would end up enhancing frequencies where the headphone is already sounding excessively sour, which would mean that the sound would be fatiguing to the ear at best, or barely usable at worst.

I knew early on that I would have to test all the headphone models, one after the other, giving each one my individual attention to see if correction using our precise mathematical operation would be appropriate or not.

I had memorized the timbre of a piece of electronic music, and that was all the help I had. The track was rich in low frequencies, distributed throughout the spectrum. A few seconds of listening through each model allowed me to understand if the correction worked with the headphone I was wearing, and whether the resulting harmonic distortion was acceptable.

In order to improve the quality of the intervention, we tried to test as many models as possible in the shortest period of time possible.

In the end, focusing on this project in a condensed period of time allowed me to achieve greater cohesion in the final product.

Once we finished screening the models that were readily available, we started purchasing the ones we were missing and visiting every recording studio we could, always looking for a new profile for our collection.

I remember a day in Alessandria when I had a tangle of twenty headphones in front of me. Most of them were discontinued models but were probably still used by many people both in professional and home studio.

We went to Viareggio, in Varese. We had had some headphones sent to us by a Sicilian friend, Alessandro Magnisi. Alessandro, being a good Sicilian, didn't stop at simply helping us with our project — he also started sending us sweets. Greta, who was in charge of logistics, spent most of her days organizing shipments, packing boxes, checking tracking numbers, and updating the work calendar.

The quest for validation during this kind of project is similar to the process of music production. There is a technical component underneath it all, but the qualitative assessment also counts for a lot. With time, I started to compile a sort of report card for each headphone, to help me remember the merits and the defects of each one.

For example, the Ollos are very comfortable headphones with good low-frequency translation. Other in-ear devices have caused us lot of headaches, as it is often difficult to get them to fit completely into the ear canal.

I've had hundreds of different models in front of me, and Antonio and I have even tried to measure headphones that were broken or in which one of the earcups didn't work. It was exhausting, time-consuming work; but in the end, I'm glad I took up the challenge. It gave me the opportunity to measure a significant number of models in a short enough span of time to be able to compare them accurately, which is a feat perhaps unique all on its own.

When John Carmak, DOOM's genius programmer, was asked the key reasons for success, he simply replied "hard work."

Every so often this phrase echoes in my mind and inspires me to do better. Often the only way to realize one's dreams is through repetition, persistence, technique, and hard work.

Many times during the testing process I felt the impulse to skip checking a particular model. Maybe this headphone was just too cheaply built or was a niche model that hardly anyone used. Who would ever notice?

But in the end, after all the time, energy, and money we'd already invested, cutting corners just didn't feel right.

I like to think that, when users put on their headphones and hear a pleasant sound, it will be partly because that model has been checked with all the passion that I have for my job.

THE REWARD

I don't remember the first time I received one hundred percent positive feedback on what we'd created. I think it was at the beginning of November 2019, just a few days before the start of the project, and I think the feedback came from Simon. As software vendors we are used to receiving comments and criticisms, especially in the preliminary stages of any implementation.

All the questions and reactions inevitably leave our interlocutor a bit undecided as to the objective evaluation of the product. Perfection does not exist, and everyone has a personal vision of the functionality that they believe each product should have.

Sienna's start was not so different from what we were used to. We received feedback saying that the stereo plan seemed very narrow; that the sound wasn't exactly enveloping; that sometimes there was an unpleasant feeling of phase shift, and so on.

But one thing was clear: there was a particular interest in the product, a different level of interest than there had been in other products that were already on the market.

Why could this be?

Perhaps the pandemic had created a new need? Everyone wanted to listen to or control a piece of music in an environment away from their traditional recording setup. Although this was a need that had certainly existed before — people wanted to be able to recreate a certain kind of sound in various environments, even on an airplane trip — it's also true that the pandemic forced us to stay home for an extended period of time. This made us think about the importance of being able to work anywhere, even in the most unexpected places, far from the comfort of our most expensive tools.

In the early days of the first lockdown, I remember the rush to set up impromptu workstations at home. My contacts were buying new converters, chairs, and especially headphones. I too bought a new chair, and I'd brought home the office computer. Most importantly of all, I had a trusty pair of Focal Clears on my new makeshift worktable.

Sienna was a product of continuous evolution. The first implementations were chock full of small errors and tiny imperfections, and so the effect was one of constant and gradual improvement.

For example, we would use a different microphone for room measurements and things would improve a little. Then, we'd measure a better room, and things would improve even more. Stefano would fix synchronization problems between samples and the sound would become richer and clearer. I would write a better linearization algorithm and artifacts that we hadn't even noticed were there would disappear.

The starting point was miles away from where we wanted to end up, but we were traveling a great distance every day. I believe it was one of Yahoo's administrators who said that we sometimes get discouraged because we feel like we haven't done enough that day. The truth is that every day is important, even the ones that seem less useful or less productive, because they are bringing us one step closer to our goal. I find this image beautiful and very motivating.

Against the backdrop of this gradual improvement, I remember three distinct phases in which I tried to engage external alpha-testers.

The first phase occurred in November, when I was very enthusiastic but still receiving guarded and sometimes controversial reactions.

The second occurred at the end of December, which is when I began to collect extremely positive and gratifying evaluations.

I had compiled a list of Italian contacts who were using the same Audeze model we'd bought, and I was very convinced of the result, I wanted to find out if they felt the same.

Jurij Ricotti and Stefano Maccarelli are two very talented producers from Rome, and their feedback was simply amazing. I was over the moon.

I stayed on the phone with them for hours, fantasizing about future implications, and about what Sienna could represent for the community of musicians and producers that would use the product. These conversations also gave me a very strong urge to do more, almost like a sort of creative bulimia.

But the best response came from Matteo Cremolini, a composer of TV soundtracks from La Spezia, who had been suffering from terrible tinnitus for a long time. In the past, I had tried to experiment with an inversion curve of his audiogram, without any particular success. But with Sienna, things went very differently. He was able to hear more, to create more balanced mixes, all while using a modestly priced headphone model made by Beyerdynamic.

Matteo was so grateful that he sent me bottles of wine the following Christmas.

The third phase started in January, when I started to once again involve the participants from the first phase, to check to see if the initial problems had truly been solved. From that moment on, things started to work consistently.

One day Carl Fath, an old friend from the period when I was involved in small-budget independent music production, came to see me. He was particularly impressed by how Sienna functioned as a sort of magnifying glass, how the software helped him to understand the flaws in his work.

With the excuse of bringing us his Grados for new profiling, he took the opportunity to chat and to try out some headphones we had just purchased: the Apple Airpods Max, which sounded particularly good. We talked all afternoon, and in the end, he was even more enthusiastic than we were! He offered us a lot of suggestions on how and what to develop in the future, and his excitement cheered us up in every way possible.

Every week since then, our office is literally invaded by people bringing support, suggestions, headphone models to sample, and great doses of encouragement. For a while we had dozens of headphones scattered around every single room, which turned the office into sort of an impromptu showroom. You could find very cheap models with unthinkably impressive performance, such as the tiny AKG 361, sitting right next to Bluetooth models whose astronomical price made them almost unobtainable.

People came to us with gifts: Francesco Campbell often showed up with bottles of Italian prosecco, and once with an expensive balanced cable for the inauguration of our newly purchased amplifier.

Cesare Marocco came to visit us with a ventricina made by Abruzzo's — a precious salami from central Italy — which we enjoyed with several bottles of wine and cheese.

I believe that this festive atmosphere was the ingredient that helped spur us on in the final stretch of our marathon. Every conversation was helpful, every suggestion contributed to our final choices.

Francesco is also the person who started the tradition of "Thursdays at the office." Every week, he would bring us a different set of expensive headphones obtained from an informal circle of enthusiasts. He'd show up with Focal Utopia or Stellia, an extremely hard-to-find Fostex model, or the top products by Audeze.

Each headphone, worth several thousand euros, was greeted by a solemn ritual of measurement — though this was interrupted by noisy aperitifs, toasts and lunches of beer and sushi.

Choose a job you love, and you won't work a single day in your life.

THE JAPANESE PARENTHESIS

Davide Burrattin is a third-party developer of libraries for one of our products called Nebula — he is also a friend who lives on the other side of the world, in Japan. Though of Italian origin, a few years ago he decided to move with his family to Osaka. He keeps us updated on the customs and traditions of his adopted country, which is so physically distant and at the same time so culturally close.

I proposed that Davide participate in the alpha-testing program in February, and after an initial evaluation, he found a balance that suited him perfectly in some of the profiles made for Bose. Noise-cancelling headphones are particularly useful in a real-world setting, as they can be worn on planes or trains, and for Davide air travel is a frequent activity. (At least, it was until the travel interruptions due to COVID, as well as the increase in ticket prices indirectly caused by the pandemic's fallout.)

One day Davide absent-mindedly told me about a Chinese headphone brand that is not very well-known in Italy but is extremely popular in Eastern countries: the HifiMAN. We phoned every friend or acquaintance we could think of, looking for an available model so we could order a pair.

After all the money that we had already spent, the idea of buying more models that might end up sitting in a box collecting dust was not particularly appealing to me. Taking into account all the headphones that we had profiled, I wondered if it made sense to insist on reviewing models that we couldn't find anywhere in the entire country. Could they really be that good?

But I decided to trust my instincts and the respect I have for Davide. When he draws attention to something, it's always for a reason and he usually ends up being right.

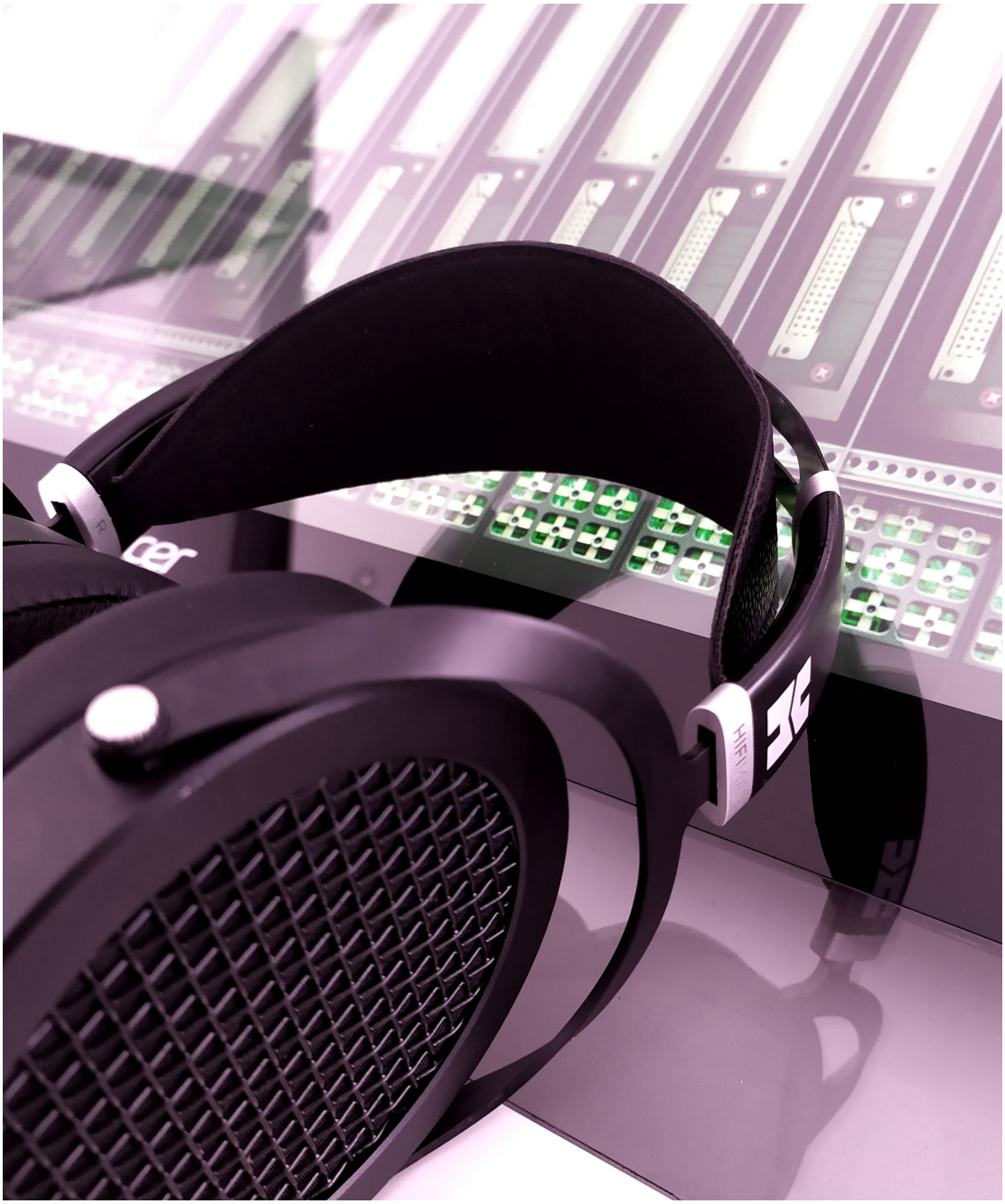
What would happen later was also a warning for the future: too often we ignore suggestions because we are caught up in our own commitments, opinions, plans, deadlines, and evaluations. The key is sometimes to listen carefully and to avoid responding impulsively.

If I hadn't stopped to reconsider, I would have missed out on discovering a fabulous company.

We bought a few HifiMAN models on Amazon, and as it turns out, they are simply fantastic headphones. As they're based on the same magneto-planar technology as Audeze, they manage to mimic the more expensive headphones perfectly, but at a fraction of the price. What's more, they're lightweight and very comfortable.

Besides our gaining new profiles for our work, HifiMAN Sundara have become Antonio's favourite headphones, as well as the model I often suggest to friends and customers who are looking for an efficient compromise between price, performance, and comfort.

Even Francesco Campbell, accustomed to top-of-the-line headphones, spent a whole afternoon trying them out. In the end he agreed that they measure up to even the best-known high-end brands.



The official beta-testing started at the end of January.

We produced a couple of descriptive instructional videos and invited our alpha-testers to join a private Facebook group. My idea was to turn it into the official Sienna Facebook group as soon as Sienna went into production.

We later invited a select group of our customers, chosen according to headphone model, and finally included our official beta testers (who usually screen our products).

The reason for this particular order was related to the predisposition of each group of users to understand the product. The first group of people were individuals who I'd personally asked to get involved, and I had already answered their every question.

The group of customers was recruited on the basis of their knowledge of headphone work: they were all people who were familiar with the competition's products and already knew a little about what to expect.

Ironically, the beta testers were the people who were the least knowledgeable about the product, as well as about using a headset as a replacement for physical speakers. But for that very reason they were also the most significant sample group. If they understood the value of Sienna without any special descriptions or assistance, it would indicate that we were on the right track.

I was also interested in their reaction for another reason: a user who had never seen the product and had never heard my explanations would probably be more objective, less conditioned to respond in any particular way. Their response would hopefully help us find weaknesses in the interface, or to find out if the software didn't work the way we expected.

The first group was enthusiastic about the product from the beginning, but it wasn't all sunshine and rainbows. Some had suggestions geared towards making Sienna more flexible and therefore more complicated. To me, the product seemed too complicated already, and I was trying to identify strategies to reduce it to an even simpler product.

At the time, Sienna was a unified plugin, set up for both headphone correction and room simulation. At first glance, it had only couple of menus and a few parameters — but there was also a button that hid an additional set of complex parameters that made the plugin particularly flexible. For example, it was possible to adjust the angle between speakers, or the volume applied during bypass. Some settings expanded the stereo panorama, others adjusted the rules applied to the pan. One parameter even adjusted the transition between pure speaker simulation and binaural listening.

The problem in product design is always the same: a flexible product generates less criticism from knowledgeable testers but potentially confuses new users. People who haven't had experience with a product like to learn how to use it intuitively, without using a manual, and still feel that they understand it completely.

From this point of view, Slate's product was perfect: it had few controls, was quite clear, and involved no complicated choices apart from room selection. Later Steven also added the possibility to choose between two profiles, depending on the width of the ear canal, and even this choice raised numerous questions.

Product experts know this simple rule: every single question has the potential to take the user further away from the purchase.

As we have seen, the differences in human geometry make it impossible to create a perfectly universal product, so some configuration parameters are necessary. But if the product involves too many options, it may seem confusing, unclear, or unsuitable. Conversely, a product with only a few choices almost inevitably seems to be the right one, provided that those choices sound right.

Of course, if it has few parameters but the result is bad, then this is also a problem.

The first criticism came from Oleg Yorshoff, one of our long-time Russian beta-testers. It was about the calibration profile for his Sennheiser, which was possibly overloaded when it came to low frequencies. Our profile had to take into account in the emulation of physical space while using our product. He still had to get used to this, as he was coming from a product with completely linear correction.

I thought of a very simple solution: dividing the plugin into a room emulation product and a headphone correction product, giving the headphone correction software an additional parameter that would optionally bring it closer to a more linear signal reproduction.

The beta testers seemed to particularly like the new solution, especially Rich Prewett, a good friend of ours from Minneapolis who had followed the development of our products from the beginning.

Rich was not accustomed to working with headphones, and so the product dedicated to simple headphone correction solved his problem of gradually approaching the new technology. (It was true that headphone correction software had been on the market for quite some time, but as he was not used to it, he needed a little time to acclimate.)

After a while, Oleg got used to the original curve without additional controls. In fact, as the linearization moves the sound towards a reference curve, it's just a matter of adaptation.

Our curve had some small advantages over the linear curve, in that there was less masking of the low frequencies, just as there is with a physical speaker system. The filter produced by the room, our ears, and our ear canal results in a sort of natural frequency boost, or positive equalization of the signal.

The second, more complex criticism came from a new tester, David Brancato, an American artist who was very knowledgeable about Slate's product.

He, like myself and a number of other users on the forum, belonged to an intermediate hearing profile and asked us to introduce an EQ control at 3.5kHz. This was in fact the frequency that created the difference between the two profiles and was where Slate's product also allowed intermediate parameterization. I was trying to cut down on the number of controls, and here came a request to add one!

I tried to pose the question to our online community as clearly as I could, writing a long post in explanation. He was probably right, the EQ control could help, but was it worth it to complicate the product further?

People responded, gradually confirming my suspicion: the control David had suggested would be a useful parameter in the quest for sonic perfection, but it was not fundamental. We concluded that adding another control would probably have fed the ranks of the undecided potential users, making the adaptation process even more complex and uncertain.

At that point, it occurred to me that it was time to stop. Just like when you're in the production phase of creating a piece of music, once all the possible arrangements have been identified you begin to reduce the elements down to the essentials. If there are too many elements involved, you don't know what to listen to anymore. If there are too many overlapping melodic lines, the piece will seem unnecessarily complex.

We started to work on two plugins with easily recognizable white user interfaces to help put some distance between us and the Slate/Realphones/DearVR Monitor group of plugins, whose GUIs (graphic user interfaces) are black.

The two plugins would have a minimal number of controls. The room plugin would work simply, with presets. We also provided the option to install a more complex and customizable product, for users who were into parameterization. But the product would have to remain true to its essence.

You can't imagine how many messages I received confirming that this was the right way to go! I remember a beautiful voice message from Stefano Maccarelli, and another from Carl Fath. According to them, simplification was indeed the correct path.

All of this brings me to the chapter on "less is more." It is often said that simplicity is the ultimate sophistication, but in school, this kind of guidance is often missing. It seems like the kind of wisdom you're more likely to learn in a professional environment.

Writing a user interface is a particularly complex job because you are creating something for someone different yourself and the other people on your team.

When we build a product, we model it according to our expectations. The origin of its functionality is tied to our specific need, or what we think the user needs.

But from a business point of view, you have to remove that personal bias even though it's a merciless operation. While you obviously have to create some sort of identity, you have to do it in a way that makes your product look symmetrical, perfect, and understandable for as many people as possible.

The truth is that today we have less time than we did a few years ago. We have less time to read a manual, less time to wonder what a control should do, and less time to spend fiddling with a parameter.

Every product must work in a direct, immediate, and almost magical way to be successful.

WHAT I REALLY LEARNED ABOUT HEADPHONES

In years past, I'd considered headphones to be an instrument with rather predictable frequency behavior. I'd tried a few different models, mostly AKG and Beyerdynamic models.

I felt like the bass wasn't particularly present, and I blamed that on the size of the driver, which was too small to reproduce low frequencies effectively. I could feel that they were there, just not in the correct quantity. Larger headphone models seemed to present fewer problems, probably caused by their increased size.

In spite of this limitation, whenever I was analysing the flaws in an edit of an audio session, I always wore headphones. This was also true for listening to the results produced by our software, as this method worked well for understanding harmonic distortion, compression, the correctness of a crossfade, or artifacts created by messy editing on a vocal track.

Later, I saw the introduction of tools such as Sonarworks, which were aimed at solving the linearization problem presented by headphones. It was obvious that there would be small differences between models, and that some software could make them sound more like each other.

Some acquaintances, friends, and even collaborators gradually pointed out the advantages of having a correction system, in the same way that it helps to have a physical system to correct speakers in a room. That way, you get used to a correct reference, and of course if everyone is working off of the same reference it's easier to work in different locations. You could go to a completely different recording studio and get used to the new listening environment quickly and easily.

I've never dealt with the problem of linearization, and I've always considered the target of the reference curve to be fairly predictable. The biggest problem is that I never thought enough about the meaning of a linear curve, nor the interference created by our physical geometry.

Let's take the example of entering an anechoic chamber: we discover that many speakers are linear; that is, they present a sufficiently flat curve throughout the range of the audible spectrum. The problem is that the sound that reaches my eardrum is not linear. My pinna, my head, even the shape of my torso will act as a filter, resulting in the deformation of that flat line.

If I'm wearing an in-ear monitor (one of those little headphones that is introduced directly into the ear canal) I will have bypassed some of these structures, like the pinna. If I send the same sound source directly to my eardrum, I will experience a completely different sensation.

Let's try to connect all the dots for a traditional set of headphones:

- The shape of some headphones means that they surround the outer ear, but still remain within the space where the sound is generated, which affects perception of the sound.

- Drivers generally have difficulty reproducing lower frequencies in a linear fashion, due to the physical limitation of headphones' size.

- The headphone essentially turns into a tiny room with a small speaker that is attached to the user's head, and all the problems of reflection that can happen in a real room can also occur when using headphones

The result is that the version of sound that reaches our eardrum is manipulated. Of course, you can change the geometry of the headphone, the materials, or some of the driver characteristics in order to reach some sort of equivalent solution. But just like building a speaker, in the end it's an art: there is no scientific or infallible solution, unless you incorporate electronic correction.

That's why measurement systems like Gras, Audio Precision, Minidsp or various dummy heads (like the Neumann model we used) reproduce an average pinna, with an average-sized ear canal to demonstrate sound's behavior in an average situation.

WARNING! These dummy ear canals are often the wrong shape! Our ear canal is shaped like a small horn, not like a cylinder.

Sometimes you can choose to buy a different cylinder to represent a different-sized ear canal, but for obvious reasons the measurement must be made for an average ear.

If you are going to build any object, you have to have some kind of reference.

In the end, the sound coming from a speaker arrives at the eardrum deformed — even if it's linear and not influenced by reflections from the walls of a room, as it would be in a fully anechoic chamber.

We can compare that sound with the measurement of an average case in order to make an approximately average correction. If we wanted to be more precise, we would have to calculate the specific HRTF of the specific user and invert it with the parameters of the measurement system in order to finally arrive at a result where a truly linear measurement arrives at that user's eardrum.

But headphones are an average system by definition, and correction systems are average systems by definition. Even the electronic corrections made by some headphones are average, unless they have sensors designed to derive the geometry of our specific pinna and eardrum.

For that reason, what we perceive as sound is an approximation — and we're not talking about a difference of a few tenths of a decibel. Our personal HRTF has values of several decibels, often placed at completely different points than other humans might have.

What is a truly linear reference curve for a headphone? Over time, many manufacturers have tried to represent a linear response that is independent of headphone and earcup geometry. It's a consequence of the sum of the physical aspects: the driver generally has a linear response, but everything around it doesn't, and the shape of your ear is inevitably one of the variables.

Some electronic correction systems try to force this situation: for example, an Apple headphone or a Bose headphone (when the noise cancellation algorithm is working), measured with Gras will show a surprisingly similar trend. Some linearization software seems to move towards a standard reference curve, which is a solid but arbitrary choice.

Other headphone manufacturers, on the other hand, have begun to make headphones more palatable to a certain user demographic by giving them back the feeling of being in a physical space. There are some studies, carried out qualitatively on a sample of individuals, that have shown a preference for a different curve that is much more pronounced in some frequencies.

Who is right? the first system will behave like the typical set of headphones to which we have all become accustomed over the years, and the second will behave in a similar way as to how we might perceive the source in a physical space.

With the addition of a curve, we throw new variables into the mix. Assuming that to correct a headphone with a response very close the psychoacoustic curve derived from the preferences of a pool of users (Harman's frequency response), the correction is usually limited in amplitude in order to avoid overly distorting the final result. This makes the conversion from one system to another extremely inaccurate and impractical.

Some websites propose Harman as a target reference curve, others propose a linear curve. If we take into consideration the result after the correction, we will have difficulties passing from one system to another in both cases due to the fact that the correction has been truncated at the origin.

Amidst all this reasoning, I want to point out an important point: it is not true that a headset that doesn't mesh well with software is inferior or wrong. It all depends on the use.

Sienna wants to reproduce a physical space. As it tries to imitate the subwoofers, it needs a very precise distribution in the low frequencies. It must allow these frequencies to be reproduced within certain practical limits of harmonic distortion. But other software may have other goals, so it's impossible to say a priori what is better or worse.

Some headphones exhibit enormous variability from one model to another, or even from one earcup to another. Some have a very slow transient response. Some can reproduce low frequencies very naturally. Some are comfortable. Some are incredibly well-built, whereas others make extensive use of fragile, easy-to-break materials. Finally, there's the question of price — a point of flexion for a certain demographic of users that can end up being an insurmountable barrier.

In the end, I learned to appreciate every headphone I was able to get my hands on as if it were a small masterpiece. When you do a job like mine, you eventually start to see all solutions as potentially valid: what changes is only the use.

In the end, this is exactly the problem. If a headphone has to be used with an algorithm that imitates speakers in a physical space, variances will emerge from its standard use both without a correction algorithm, as well as with a correction algorithm declared as linear (which is the case of Sonarworks).

In the end, I feel extremely positive about correction and emulation software. I am so grateful to be living in this moment in history, surrounded by incredible tools built by exceptional people.

Slate's system is impressive, and so is Sonarworks' system, but the reason I set out to build a new product is because there was still a hole in the market that needed to be filled. Sonarworks tried to build a linear, or classic reference system. Slate's design was aimed at creating a solution for its headphones only.

Realphones is an incredible product but is currently limited to a particular implementation of HRTF, among other specific parameters.

While it's true that we approached the problem with the intention of simply creating a commercial product, in the end we built something that was really missing from world, and that we personally felt was needed on the market.

And the headphones that make this product functional are by no means the best or the worst available.

For example, AKG over-the-ear headphones usually perform poorly in terms of low-frequency reproduction when compared to other brands. This is not to say that AKG is a bad manufacturer: their headphones are incredible, the design is often very nice, they are comfortable and solidly built, and within the same product line you can find inexpensive K361s that work very well. Over time, we have found that AKG over-the-ear is one of the better solutions available. They have a fast transient response, which often allows you to hear details accurately.

Sennheiser makes headphones that work reasonably well: they're comfortable and sufficiently linear. They're not amazing headphones, but they do their job. If I had to point to a headphone with decent output for our algorithm, a low price, and very good build quality, I'd point to the mid-range Sennheiser models. However, it's a headphone that distorts easily in the low frequencies, so if it's the reproduction of the sub that is of the most interest there are probably more suitable brands.

Apple makes some incredible products. Its Airpods Max are insanely good quality: they offer low harmonic distortion, an extremely fast transient response, and high sensitivity. The other Airpods (the non-"Max" models) have the problem of not enough sensitivity. In all of their models I hear a lot of digital artifacts, so I'm not sure they would be the ideal tool for a long mastering session, but it's worth owning a pair if only to compare how the end user might perceive a mix when listening through them. This problem of digital artifacts applies not only to Apple, but shows up in all products with a non-wired transmission.

Audio-Technica produces models with extremely variable output: some models tend to have generous bass response and even rumble in the low end, while others are more balanced. Others tend to distort easily, even in the high frequency range. In general, I was forced to limit the corrective range of intervention for sub frequencies when using their headphones, but they are usually inexpensive, comfortable, and well-built. I really appreciate they're detail and transient response. With any luck, I'll have the opportunity to experiment with the more expensive models, which I haven't yet had the chance to profile.

Ollo Audio is an up-and-coming manufacturer that offers a range of nice, reasonably priced products. Low frequencies are not their strong suit, but their sound is extremely pleasant and round. They generally offer good definition, and one of their unique selling points is that they can be worn for an entire day without ever getting tired — as Gianni Bini explained to me one day at lunch, while we were eating sushi.

Slate Audio has produced only one headphone model, but it is incredible. The low-frequency definition is exceptional for its price, and it has one of the fastest transient responses I've ever heard. It's definitely my favourite headphone when it comes to understanding harmonic distortion. The construction feels a bit fragile, but I haven't encountered any particular problems while using a pair. The model is lightweight and comfortable and is probably one of the best closed-back headphones available. Its sensitivity is also exceptional.

Sony is a headphone brand that I don't particularly adore, but I recognize the reasons for its widespread popularity. It's a headphone that tends toward a metallic timbre and sometimes has limitations in the low end, but generally offers excellent definition. It's an ideal headphone for hearing detail during the editing of a vocal section, for example, and their models are usually inexpensive.

Focal is a brand that I love. We have a lot of their speakers in the office, and it was also the first high-priced set of headphones that we ever purchased. It has its limitations in terms of delivering low frequencies, but its definition is unparalleled, and its midrange and treble responses are absolutely surgical. Listening through the Focals really allows you to focus on the music. Their build quality is crazy (in good way), as is their choice of materials. The headphones are lightweight and comfortable. The lower cost models also offer very high performance in the low frequencies, but at the expense of some loss of detail.

Audeze is the headphone I would take if I had to spend the rest of my life on a desert island. Built around magneto-planar technology, these are the ideal headphones for representing low frequencies as if they were being reproduced by a high-quality speaker. In general, they have a somewhat slow transient response, but as this can also be said of physical speakers, the auditory illusion they create is perfect. Only the top-end models are comfortable and lightweight; the others are well-built, but bulky and heavy. They provide the perfect room simulation, provided you work in fairly short sessions.

I talked about **HiFiMAN** in the chapter "The Japanese Parenthesis." They, too, produce a number of models based on magneto-planar technology and are often an ideal, less expensive alternative to buying an Audeze model. However, their catalogue also features high-end products, some with astronomical price tags. Someday, I'd like to profile the HiFiMAN models that we didn't have opportunity to examine during this project, to find out what level of sound quality they manage to reach.

Beyerdynamic is an extremely popular brand. They are comfortable and lightweight, and their design is often imitated by other manufacturers. However, they are not the ideal choice for reproducing the range of a subwoofer, and sometimes suffer from quite a bit of harmonic distortion even in the highest frequencies. Physical space emulation algorithms can work very well with them, but it's not their ideal terrain. In spite of these limitations, they have been my favourite headphones for a long time and are also the most popular headphones in our office. You can wear them for hours without feeling fatigued by their weight; and in spite or perhaps because they're missing some detail, the listening experience is pleasant and never harsh.

If the Apple Airpods Max didn't exist, I would probably point to Bose headphones as my brand of choice for noise-corrected listening.

They do have a few minor limitations: a slightly intubated sound, and sometimes the treble reproduction is a little laboured, but other than that they're a great product. They can be used anywhere, including on a train, a plane, or even the subway. For years we've used a Bose model at trade shows to show off our algorithms in the middle of a room filled with ambient noise and crowded with people.

Bower & Wilkins is probably my favourite speaker manufacturer, and their headphones are also particularly well-constructed.

In some models, the noise cancellation algorithm seems to create a slight dynamic compression.

Grado makes a wide variety of models which offer very diverse results. They are exceptionally well-built headphones, even if they are not particularly comfortable. In all of their models, the ability to reproduce detail is phenomenal and the transient response is very fast. The reproduction of low frequencies is often their weak point, but some of their more expensive models manage to remedy even this detail.

Presonus is a brand that releases products of incredible quality at extremely favourable prices. All the models I've had the opportunity to measure have displayed exceptional performance, as I described in the chapter "The Journey." Their weak point tends to be in their transient response, as some details can get masked. But considering their low retail price, I wouldn't ask too many questions.

Shure is a very popular brand in Italy, with discrete products at often very affordable prices. Their weak point is the accurate reproduction of low frequencies.

Superlux is the king of budget headphones. In spite of their modest prices, their products look very professional and offer good build quality. Their weak point is their high frequencies, which tend to be strident to the point of being annoying. Correction algorithms can take this into account, but this solution has its limitations. In fact, the variability in their products' production forces little selective interventions, in order not to run the risk of worsening the situation. The low frequencies are present, but only slightly, even after corrections. The possible interventions are limited by their tendency to create harmonic distortion.

Ultrasone headphones represented a bit of a challenge for our correction algorithm — you could say that the Sienna linearization process actually began with them. Generally, they are headphones of good quality, comfortabl, and detailed enough, though perhaps a bit laboured in the treble range. I don't particularly like their sub-range reproduction, but the level of harmonic distortion is acceptable.


HD 650

THE CREATIVE PROCESS

I often wonder about the creative process: what is it exactly? How does it work? While searching for a suitable catchphrase as an introduction to this story, I came across an aphorism by Steve Jobs that I thought was absolutely perfect.

I consider myself to be an average person, with an average level of creativity and average problem-solving skills. I was an inconsistent student at school, with lots of ups and downs throughout my academic career. I often forgot things and was forced to study them all over again.

I'm particularly interested in describing what the creative process is like for someone who belongs to the world of the average, because that's exactly my case. A few years ago, while reading a book that told the story of a group of computer programmers, I came across a sentence that went something like this: "I'm not the best in the world, but I'm well-rounded, I adapt easily."

Each one of us — even the least likely, or least obvious — possesses extraordinary characteristics of which we are often unaware. Sometimes our talents are hidden deep inside, and we struggle to uncover our own strengths.

For example, Nirmal Puja — the man who broke global mountaineering records by climbing fourteen mountains, each one over 8000 meters tall, in just six months — would never have realized that he could withstand critically low oxygen levels at extremely high altitudes if he had never experienced that situation. In his own words, Puja was a fairly average runner at sea level. His exceptional quality was that he turned out to be just as fast at altitudes where most people would have to pause for a minute or more just to catch their breath after taking a mere seven steps.

Some of us who are mediocre at traditional intellectual tasks (such as mathematical analysis) can be very good at connecting elements of disparate and distant domains. Over the years, my work has allowed me to realize that — while I may be a mediocre music producer — I possess an off-the-charts ability when it comes to accurately and consistently profiling a large number of headphones in less than a month. Not only that, but I can do it while also balancing dozens of other activities, such as business management, programming, debugging software, and public relations — and with almost no outside help. Some of the audio profiles I did by ear were later found to be accurate when the correct measurement tool arrived.

This goes to show that if we were to reduce everything to pure technical ability, we wouldn't get very far. The world is full of exceptional people with uncommon technical and physical characteristics, but only a limited set of those capabilities make it to the end of the creative process, where is where you end up with a product in hand.

Let's face it, product adoption by the public is just the last chapter of the story. It might validate what you've created, but it's the last link in a long chain. During the creation phase we often talk about getting to the "bottom" of a product, but that does necessarily include getting a consensus from its audience.

At the moment in which I am writing these lines, it's not possible for me to know whether or not Sienna will be a commercially successful product. But that doesn't really matter: Sienna has already passed its first test. All our alpha- and beta-testers have admitted its superiority over other commercial solutions, which is almost unheard of.

We are just a small software company, and I had to create and study most of Sienna's elements by myself. When someone, even just one tester, comes to the conclusion that our software improves the performance of a headset manufactured by, for example, Apple — the most powerful company on the planet — it's clear that our first goal has already been achieved.

So, what made the creative process possible in this case? How did we get where we are today?

As I described in the first few chapters, we embarked on this small yet titanic task out of recklessness, as well as a lack of knowledge as to what we were getting ourselves into. Small, because it was only a minor activity in the vast universe of scientific trials and results; but also titanic, as we were just a few people with limited means, trying to achieve a difficult thing.

If we'd understood the magnitude of the problem, we would have dedicated ourselves to the construction of an intermediate product; for example, one that was limited to the overall correction of headphones only. However, the result would probably not have been the same.

Our algorithm is particularly significant when it is compared with another algorithm that has critical issues in certain parts of the audio spectrum. If I take Sonarworks' correction and apply it to Sienna's physical space emulation system, the result is not optimal at all. As explained in an earlier chapter, if my algorithm involves emphasizing low frequencies and the correction algorithm did not take into account the limitations of the specific headphones I'm using, I end up with an imperfect result.

I wanted to describe in detail the creative process in order to precisely to show the steps we took:

- Initially, general ignorance and lack of understanding of the magnitude of the specific problem was our driving force. Sometimes we need to underestimate a problem, to take up a challenge we may not understand, and not allow ourselves to get discouraged before we begin to solve it.

- Problems appeared one by one, so we dedicated ourselves to finding a single solution for a particular issue before moving on to the next one. We used all the means at our disposal: internet documentation, scientific papers, even blogs, combined with our own empirical experimentation. We studied. We tried. We formulated new hypotheses and arrived at new solutions.

- Each problem we solved led to other problems, but also to new encouragement.

- The model of what we were looking for already existed, which showed us that what we were trying to do was feasible. It helped us not to lose our way.

- Sometimes we found new solutions simply by connecting distant dots that had been right in front of us for days, or even months. Think of it like climbing a mountain, and suddenly finding a new path towards the summit. It's a process of research and connection, of discovery.

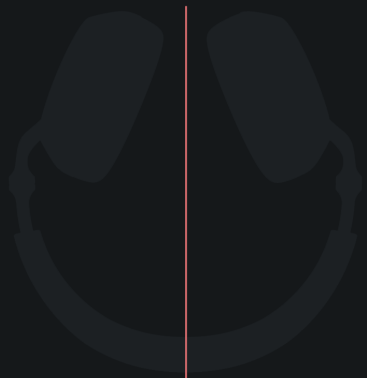
- While it's true that scientific progress needs a scientific approach — that is, a mathematical and elegant solution to a problem — it is also true that a practical application is needed when your goal is to build a working tool. The user of our products does not need a mathematically perfect solution to his problems but needs a better solution than the one he or she already has.

As Jobs pointed out, sometimes we feel like impostors when we're in the process of creation. But in the end, we didn't do anything extraordinary, we just connected the dots — maybe in a way that no one else ever had before.

That's what progress is all about.

ACUSTICA

THE HEADPHONE WHISPERER



SIENNA