BREAKING THE SOUND OF SILENCE

Voice Analysis, Criminal Investigations and Dolphin Communications

Listen carefully. Even in the quiet of the night, you probably notice a car in the distance, a clock, a fluorescent light, a refrigerator, all making their presence felt. The world is full of sounds. The refined senses and technology of researchers enable them to reconstruct scenes that ought to be invisible from these barely audible snippets of sound. Or try to understand voices that are not voices, made by animals that work at communicating by secret means that leave us humans far behind. Let us slip below the floor of the sounds we are familiar with and go in search of a much deeper world.

A World Changed by Voice Recognition
The 16-Second Mystery

"Please listen to this. I wonder what will draw everyone's attention." The tape plays. The sound of a push button phone being dialed is heard. A woman's voice says "Who is it?". There is no reply. Just the sound of buttons being pushed.
"It's a silent call. A prank call. So what do we deduce from this?"
None of us had an answer to this question being asked by Dr. Matsumi Suzuki, Director of the Japan Sound Research Institute. After all, we had only heard the sound of the buttons being pushed...
"Right at the beginning, you can hear the sound of a 10 yen coin dropping. That tells us the caller is using a public phone. After the phone is put down, you can hear the pulse tone and by analyzing that, you can generally figure out what exchange the call came through. Other than that, you can hear the sound of two vehicles, at 20 second intervals. From the sound of the engine, they are probably trucks. Since frequencies are altered by the Doppler effect, that enables you to calculate the speed at which they were running. The recording was made at 10 p.m. and the fact that two trucks went by within 20 seconds suggests that it is a fairly busy street. By putting all these conjectures together, the idea is to stop whoever is making these prank calls."
It's like watching a famous chef wielding his knife. Even with just basic listening, an enormous gap is evident between us and Dr. Suzuki.
"This is how I make my living, so there is definitely a professional's sensibility at play."
Working as a "Crime Detective for Sound", Dr. Suzuki has provided important information that was used both to help investigations and as evidence in court following major crimes such as Benito Aquino's assassination, the Korean Airlines bombing and the Aum Shinrikyo incident. It was probably a bother for him but he even received a note from the "Man of 21 Masks" who perpetrated the Glico/Morinaga chocolate poisonings, challenging him "Hey Suzuki, the Sound Man, do a good job of the scientific analysis!". He is indeed seen as a key figure for criminal investigations.
When a young working woman was kidnapped in Kofu in 1993, he plotted the following route and traced the telephone from which the threatening call had been made.
"Since long distance calls use microwave circuits, filters are used to cut low frequencies. As the threatening call had these frequencies, I concluded that the call had been made locally."
"Starting with the different frequency characteristics of new and old exchanges, I was able to pinpoint the area from which the call had been made."
"Using the fact that each telephone within an area emits slightly different noise levels, we were able to pinpoint the unit."

MATSUMI SUZUKI
Director of the Japan Sound Research Institute. A graduate of the Faculty of Science and Engineering of Kinki University. Ph.D. from Adam Smith University. Also studied at the Sorbonne University and Gifu College of Dentistry. Has held posts at the Scientific Investigation Research Center of the National Police Agency and the Science and Technology Agency. He studied at the KAY Sound Research Center of the FBI in the United States. He also receives many requests other than to investigate incidents and has been involved in producing a CD of dinosaur sounds. His writings include: "Crime Detective for Sound."

The representative piece of equipment used by Dr. Suzuki is a sonograph (a unit to analyze frequencies). It displays as light and shadow whether a sound is high or low, strong or weak, as well as time variations. Based on this, he can make subtle differentiations between the way a voice vibrates, the way a mouth moves, the speed at which the tongue moves and the actual size of the inside of the mouth. It is possible to differentiate people quite precisely with just their voice patterns, and even possible to determine their age, height and emotional state.

Voice patterns made an impact on the world in 1983, at the time of the Aquino incident. The first announcement made by the Philippine Government stated that as security troops were escorting Mr. Aquino down the steps of the plane upon his arrival, he was shot after taking two or three steps. The communist element who was identified as the killer was executed on the spot.

In fact, Mr. Aquino was accompanied by camera crews from ABC of the United States and TBS of Japan. They were held back inside the plane and unable to film the assassination, but their cameras were running and recorded all the sounds.

Dr. Suzuki made us listen to that tape. Against a lot of background noise, we could hear several impatient voices. Then the sound of a gun. A woman screaming. Then, three more gunshots. It's a vivid 16 second record.

Dr. Suzuki analyzed each of the voices from the voice patterns, and clarified the conversations which had been inaudible due to the background noise. "Acona" (He's here). "Etona" (I'll do it). "Pushila" (Shoot!). When he then went on to analyze the voice patterns of Aquino's military escorts, they were one and the same.

But that wasn't the end of it. The Philippine Government made no effort to use the results of Dr. Suzuki's analysis.

So, Dr. Suzuki embarked on an even more in-depth analysis. "The fact that the recording had been made by two television cameras meant that you could make use of the stereophonic effect and measure the source of the sounds. But to do this, the sounds from the two microphones has to be modified to the same recording level. Luckily, we had the necessary materials. Listen to the tape once more. ... There! can you hear it? That sound like static? It's the sound of the engine of the CAAC plane that Mr. Aquino flew back. Since the sound of the engine is coming from a distance, you can probably assume that the level of the sound reaching both microphones is about the same. Using that as a baseline, we evened out the two recording levels, and by putting them through computer simulation, worked out where the gun was fired from."

The next thing was footsteps. On the tape, the sound of Mr. Aquino going down the steps has been properly recorded at 0.8 second intervals. But even that stopped at the 11th step and was replaced by interference.

"But if you think of him continuing to go down at 0.8 second intervals, even after that, unless he jumped off half way down, there is still no way he could have gone down all 19 steps by the time the gun was fired. The announcement by the Philippine Government that he was shot after he had descended and taken a few steps made absolutely no sense."

The deciding factor was the sound of the gunshot. The Philippine Government had announced that a .357 Magnum had been used but the sound pattern analysis of the gunshot revealed it to have been an exact replica of the sound of Colt 45. That is the gun carried by the Philippine armed forces.

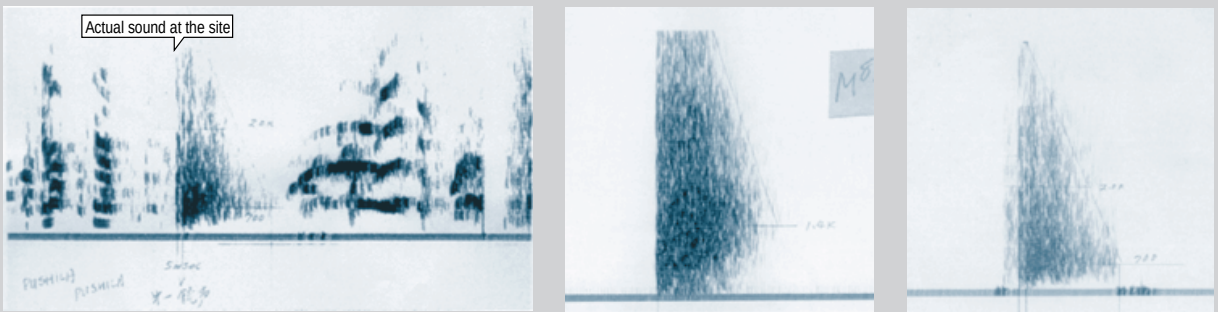
The Complicated Parameters of the Audio-Visual
Machines Can't See through Habits

"People often make the mistake of thinking that if you put something to a machine, you will get a reply. But the fact is that it all starts with listening with your own ears. First you listen to the whole, then if it's the sound of a car, you concentrate just on the car, and you do it again and again and again."

At the time of the Aquino incident which is related above, there were 50 to 60 items that needed concentrating on. Since it's normal to listen to each of them over 10 times, the total count easily topped 500 times. "Since there are more than 100 possible types of computer analysis, you have to listen to each item very carefully, and decide which would be the best method for that. So, it's quite a process until you can line up all the samples you need for analysis."

n **The Sounds of the Gun with Which Mr. Aquino Was Killed and a Magnum .357**

The on-site recording (shown on the left) shows that the sound was not, as had been claimed, from a Magnum .357 (shown in the middle) but from a Colt 45 (shown on the right).



When you hear this kind of thing, you feel that the audio visual senses of people are pretty amazing.

"It goes back a long way. For example, once a person gets used to it, they can mask background noise and only extract the signals they want to listen to. If you listen to the same tape on a machine, all you will hear is background noise. Machines are bad at going beyond irrelevant things such as white noise. At our research institute, we have finally developed software with that ability, but it requires an enormous number of computations."

With this software, would it be possible for anyone to hear "inaudible sounds"?

Just as Dr. Suzuki replied "To a certain extent, once you get the hang of it", one of the researchers who had been playing back tapes for us for a while (Dr. Hajime Suzuki, Doctor in Science, and in fact, Dr. Suzuki's son) opened his mouth to say "but, since there are different kinds of static".

"In order to eliminate the static, you must first grasp the characteristics of the static. That analysis depends on the type of background noise, plus you have to do accumulative analysis while effecting barely discernible changes in the parameters."

It is the skill of the engineer which determines whether the equipment's capabilities can be used to the utmost. When will there be a machine which will automatically recognize anyone and any kind of voice? Dr. Suzuki who is so well aware of the difference between individual voice patterns and the depth of voices, replies: "There are still major barriers to this."

"For example, if you compare the voice patterns when someone is saying the five vowels separately, and then saying "a, i, u, e, o" in a natural progression, it comes across as completely different. When language flows naturally, your mouth is already preparing to say the "i" while you are still saying the "a". So, you are influenced by the next sound you are going to make. As a result, even if you record the sounds "a", "i", "u" into a computer, it still cannot understand them when they are a succession of sounds."

So would it work to record them as words?

With each person's voice having individual characteristics, the voice patterns would not be the same. So, how does a person's audio-visual sense process this?

"I think that we probably suddenly change parameters depending on who we are talking to and the conversation, deciding that certain words end in certain places. For example, if you are speaking from someone from the Tohoku area of Japan, you will adjust your own parameters to fit in with this."

That kind of flexibility is lacking in computers, so that the simple fact of expanding the circle of tolerance is linked to errors.

"Still, you can record your own voice characteristics into a card, and when you are using a computer, you can just slip it in and reset the parameters to your own. That is probably the quickest way to do it."

After the conversation had moved away from criminal investigations, the room's atmosphere was softened with a discussion of recreating the voices of historical figures, and analyzing the sounds made by dolphins and whales. "That kind of research is pleasant since you don't risk your life." says Dr. Suzuki. His son, with his usual precise timing, makes us listen to the songs of whales.

We all relaxed.

n **Voice Patterns When Vowels A,I,U,E,O Are Said Separately**

n **Voice Patterns When Vowels A,I,U,E,O Are Said Together**

Creatures Who Command High Frequencies

The Whispering Language of the Sea

Even though the cherry trees are in bloom, the beginning of May is still cold in Hokkaido. But the lecture room at the Muroran Industrial University is bathed in a candid mood brought about by visiting students from a number of different countries. It's an international presentation and research meeting. Assistant Professor Takashi Uozumi of the Information Engineering Faculty is giving a lecture on "The Study of

Methods for Listening to Ultrasonic Waves Emitted by Whales."

The documents below were part of that presentation. "The porpoise avoids several bandwidths which can be heard by its enemy, the killer whale, and emits sounds on both sides of these. It continues to have secret conversations while listening to the sound of the killer whale."

You shouldn't feel sorry for the killer whale. In the case of human beings, however acute their hearing, the highest frequency they can hear as sound is at the most 20 kHz. Compared to the white whale which can hear frequencies up to 300 kHz, it's a though we were permanently wearing ear plugs.

Even so, the general impression of dolphin study = an endless expanse of blue sea plus beautiful sandy beaches with palm trees = is one of the tropics, so what's this in Hokkaido?

After his lecture, Professor Uozumi started by voicing these arguments.

"Since the Uchiura volcanic bay off Muroran's Chikyu Cape is where cold and warm currents intersect, it is very rich in plankton and attracts many fish. It is an attractive environment for fish-eating dolphins. And since dolphins are mammals with no gills, mothers need to prop up babies that cannot swim adequately and make them breathe. The volcanic bay with few waves is well-suited to raising young. Every year in late May a few thousands of dolphins come, give birth and raise their young."

This means that it is possible to observe various types of communication between dolphins that are calm and leading their life. Furthermore, the sea is dark up north. Since dolphins produce more sounds when they are blindfolded, the fact that the sea up there is not light is good for the study of their sounds.

As can be seen in the graph below, there are many different types of dolphin sound. Having been warned by Professor Uozumi that: "they have been classified for expediency but in fact, I think they use them in a more complicated fashion", he made us listen to the sounds that had been collected.

The first was a click. The numbing-like sound of the Pacific white-sided dolphin. It's a sound that shifts from high to low, just like a child imitating the flight of a hand-held model plane. (Graph 1)

"The cries towards the end are using changing frequencies to probe the size, shape and substance of things. The higher the frequency (short wavelength), the better it is reflected off smaller objects, and the type of reflection depends on the substance."

By the way, it is said that the bottlenose dolphin can differentiate, with a 100% accuracy rate, between aluminum, iron and coral, using these powers.

"I think that this ability to recognize objects using ultrasound could be used to make equipment for helping visually impaired people."

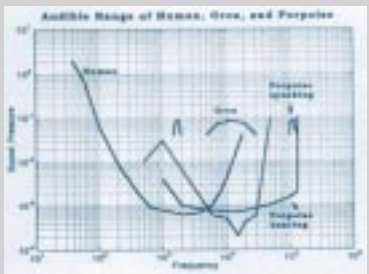
Next came the whistles of the pilot whale. The two kinds of whistle were quite different from one another.

"If you look at a sonogram, there is a dark section which is a nice continuous pattern. This pattern is close to a human voice."

It is this that leads us to think that it is used for communication.

"The variety of whistles is limited, however, and I think that clicks are also used for conversation. The high frequency clicks quickly fade and don't have much reach so I have a feeling they are used for one on one conversations. In addition, there have also been research papers on continuous emission of accumulated ultrasonic beep-like signals, in a way similar to packet communications of the kind used in the Internet. This is probably in order to improve the S/N (Signal/Noise) ratio. If they are indeed using communications technology that is similar to what human beings have come up with, that is interesting, isn't it?"

n **The Hearing Range of Human Beings, Killer Whales and Porpoises**



TAKASHI UOZUMI

Assistant Professor at the Faculty of Information Engineering of Muroran Industrial University. Ph.D. in Engineering. Graduate of the Electrical Engineering Faculty of Muroran Industrial University. Completed the doctoral research course of Hokkaido University's Engineering Faculty. Studied at the Saskatchewan Provincial University in Canada, and developed software for handicapped children at the National Special Education Research Center. Started his research into dolphins and whales two years ago. His principal occupation is to process live signals using artificial intelligence. He is the author of "Controlling the Heart and Blood Vessels."

Technology to makes ultrasound audible

Listening to Voices that Aren't Voices

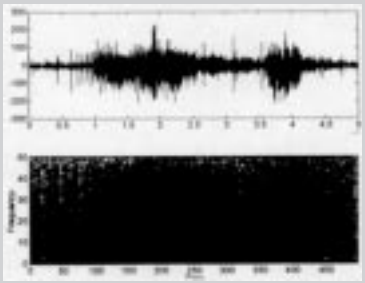
Let's take a look at how high frequency calls that are inaudible to the human ear are studied.

First of all, an underwater microphone that can be used up to about 100 kHz is brought aboard a ship, and the cries of dolphins and whales in the ocean are recorded. The next step is to put the sounds through an A/D (Analog/Digital) conversion, and at that stage they are input into a personal computer and

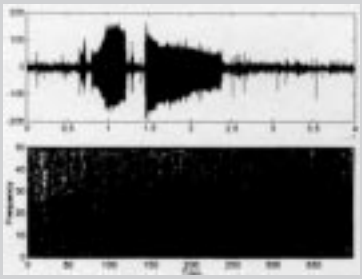
n **Varieties and Uses of Dolphin Sounds**

Click	0.001 to 0.01 seconds, 15 to 300 kHz. fi Echo Location (Recognizing the position, shape and substance of things)
Whistle	Continuous low frequencies. fi Communication
Bark (Moan)	High pitched. fi Anger, Love Calls

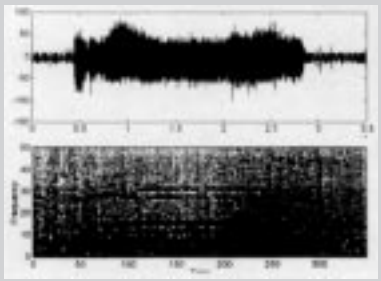
n **Waveform and Sonograms of Dolphin Cries**



l **Graph 1 : Click**



l **Graph 2: Whistle - Type 1**



l **Graph 3: Whistle - Type 2**

shown as graphs, with the sonogram displayed using a software application. In order to make an accurate data rendering of the sound information, it is necessary to implement sampling using more than twice the frequency.
As a result, the data is enormous.
A 5-minute, 12-bit sampling at 300 kHz is 180 megabytes of data.
Professor Uozumi has also embarked on research that goes one step further.
"I thought that rather than just seeing the ultrasonic voices on sonograms, it would be easier to grasp them if we could actually hear them."
By making them audible, our deep-set hearing mechanism could perhaps give some important hints for voice recognition.
"If you put it simply, it's enough to run the tape recorder at a much slower speed. But if you change frequency as though you were just trying to pull the wavelength into a straight line, you get an indistinct, sluggish sound. At that stage, you need to bring it to a slightly higher frequency that is better suited to the human ear and at the same time reduce the parts that are difficult to hear."
By doing this, the sound of dolphins comes leaping out of what had appeared to be a silent sea.
"So, by using this method, we are now working on an analysis model that enables us to replicate dolphin sounds using little data. If, based on a quantity of data that can be shared on the Internet we could show sound information that more or less matches the original frequency, it would make it much easier to have discussions from scholars all over the world. Also, the data from dolphins that live in the northern seas is apparently quite rare in the world."
Next summer, the "International Dolphin & Whale Conference" will take place in Japan for the first time, at Muroran. At about the time this magazine is published, Professor Uozumi will probably be regularly out at sea, gathering data for his presentation at the conference.
The morning after the interviews, we went to Muroran harbor. The port which was once a gateway for this steel-manufacturing town, is now quiet. The railway lines are covered in weeds. On the far shore, the reddish brown roofs on the factories are like a landscape full of memories.
But if you listen carefully, you can hear the hum of the expectant voices of the researchers who have gathered here from all over the world.

n **The Synchronized Swimming of the Pacific White-Sides Dolphin**
Photo: Tadao Furuya, Provided by: Volcanic Bay Marine Life Observation Society