

Speed of Sound

Atau Tanaka

*Published in **Machine Times**, NAI/V2_Organisation, Rotterdam, 2000.*

Music and sound are merging as art forms, with less to distinguish one from the other than concerns they share of material and structure. If in music the 20th century was preoccupied with freeing itself from the limitations of tonal harmony (1), indications are that in the 21st century focus lies already in leaving pitched vocabulary all together. If the art of sound arose in the early 20th century as a vision to grasp, capture, and manipulate sonic phenomena (2, 3), we begin this new century extending and applying this vision to social domains (4). In both cases, focus has been concentrated on the pure elements that constitute sound. The primary and fundamental component of music and sound is time, for sound is immaterial – it can only be described as functions of time, and exists solely as perturbations of energy through a medium and across time. Time in sound operates on a multitude of levels – from the macro-level of large scale musical structures, through micro-level details inside the sonic event, to nano-level temporal structures inside our physiology.

The speed of sound is not fixed, but changes as it traverses different media – air, water, electrical circuitry, digital logic, the human body. The speed of sound as we know it, approximately 340m/sec (5), is the speed of transmission of sound as pulsations of air. This figure is not a constant but varies as a function of humidity and temperature (6). In water, a medium with different transmission coefficients, sound travels at 1480m/sec. The speed of sound changes completely, or becomes abstracted in the machine as acoustical energy is transformed into electrical signals. The existence of sound and audio signals across these different planes forcibly has an effect on the way aural phenomena are assimilated by human perceptive mechanisms.

Material

At the macro temporal level, sound exists to create an invisible architecture marking the rites of daily life. The Divine Office or Canonical Hours of the medieval church not only celebrated devotion but tied it to demarcation of time through the use of specific musical forms. In simplified form, church bells broadcast out to the populace to mark the hour. The Church thus consolidated power over its congregation putting sound and music to use as keepers of time. In a secular realm, the rooster marked the start of the day, and remains a strong image describing time periodicity at the quotidian level. In machine times, these sonic elements begin to be manipulated – the mechanical cuckoo announces the hour, and the tick-tock marks the second – we begin to focus in on smaller units of time. The alarm clock indicates a desired time in sound, exercising a wish to control our relationship with time.

These examples demonstrate how time can be turned into a commodity and tool of power (7) through the service of sound. Occurrence and regular periodic repetition of sound ground our sensation of the passage of time. Once established, the converse becomes possible –exploiting time to commoditize sound in the form of music. It is through structures of time that musical form is defined – be they religious masses, symphonies, or

pop songs. Formulas of time have been meticulously refined to stimulate the appreciation of music, often for the goal of capital gain. These have been closely tied to the mediums and materials of music delivery. Court sponsored composers of the Enlightenment were careful in their symphonies not to push the patience of their audiences in the concert hall. The ideal length of a pop song is linked not just to attention span, but historically to the maximum time fitting on a 7" 45rpm vinyl disc. The Beatles harnessed the longer time format of the 12" LP to create narrative structures of the "concept album". The even longer capacity of CD's now surpass our effective attention span. Artists exploit this to create works shorter than the maximum time, then adding bonus tracks or surprises after long durations of silence. Raves as all-night/all-day techno music events create a time continuum, expanding time to a near state of suspension at the macro-level, albeit propelled at the micro-level by clocking the pulse to the human physiological rhythm.

Manipulation of time, then, is not just part of the technique of composing music, but is a tactical means for commercial entities to make a product out of music. The ultimate manipulation of time is to freeze the progress of time altogether. This can be in the form of fixed media such as scores or recordings, abstractions in the creation of a personality driven star system, or documentation in the context of historical perspective. By removal, or capture, of the time element, these forces seek to establish an existence for music as a material commodity. Otherwise, time in its natural state, and thus music as a temporal form, leaves no trace.

As the church was resourceful in using sound to mark the passage of the day, it was equally clever in utilizing qualities of sound to demarcate space to assure a position of power. The reverberant qualities of acoustic inside a church is perceived by our modern conscious as something beautiful – creating a sense of tranquility relative to the stressful bustle of the city. However, for the average human conscious of five hundred years ago, the effect must have been quite different (8). The same reverberation was at once more contained than the open air of the field, but more open than the narrow streets of the village. To the villager entering the church, the acoustic must have invoked an aura of scale, of otherworldly power. The sound reflections help to underscore the large cross-shape architecture of the church. The church, it can be said, mastered subliminal special effects long before broadcast media. The use of temporal properties of sound was a key strategic element.

Perception

If such use of sound diffusion was effective, it is because the human physiology is sensitive to, or in fact depends on, these cues. Reverberation is a cue that helps us to orient ourselves (or in the case of the church, disorient ourselves) in the space that surrounds us. Simple surfaces – a building across the street, a mountain cliff hundreds of meters across a valley – give us distinct echoes. Parallel surfaces – the walls of a staircase – give us repeating echoes that can merge into a pinging buzz. The complex dimensions of a church meld echoes together into smooth reverberation. The reflective acoustical properties of these surfaces, and the dispersion time of sound between them, govern these effects. Beside the readily demonstrated examples of mountain echoes and pinging staircases, the human perceptive systems are sensitive to far more subtle effects over

shorter distances and lower amplitudes (9). In the same way that echoes tell us how far away a mountain is, shorter echoes can help orient us in our immediate space. The time that it takes a sound to come out of our mouth, reflect off the ground and come back up to our ears, remind us how tall we are. The modification of this effect as a function of the surface on which we are standing – an asphalt street, a grass field - change this, as absorptive materials dampen reflections. The acoustics of a snow-covered town gives us a “quiet” peaceful feeling, and perhaps even a floating feeling as we have fewer cues indicating our height. Human aural perception is sensitive to these subtle changes not just at short distance, but at larger distance as well. A clear sunny day feels more open than a cloudy day not just because we appreciate the sun, but because there are fewer reflections from above – clouds reflect sound, and lower clouds mean shorter reflection times than higher clouds. Unaware to us, the sounds of our environment are reverberated by the sky.

In discussing these subtleties of acoustics, we are dealing primarily with sensitivity to effects of low amplitude. Human auditory perception is equally sensitive to subtleties at the micro-level in time. The principal mechanism for distinguishing the location of sound in a horizontal plane is the difference of time of arrival of a sound in the two ears (10). The width of our head is the extra distance that a sound on one side of the head must travel to reach the ear on the other side of the head. It is a distance small enough that the amplitude of the sound does not diminish so much, but is enough of a distance for the auditory perceptual system to distinguish time of arrival. At a speed of sound of 340m/s, the difference in time of arrival over a distance of 20cm is 59 μ s

Human auditory perception has its limits as well – the minimum time in between sonic events to be distinguished separately is on the order of 20 milliseconds (11). Inside this limit, sounds begin to blur together to create first a buzzing then a continuum. Although this may be the limit to distinguishing sounds as separate events, human hearing is extremely sensitive to timing quality of events spaces in time. Percussionists create musical sonic events that are distinct, typically separated by hundreds of milliseconds. The accuracy with which they articulate any given event, however, is on the order of a millisecond. Variations at this level create different musical feelings for the same notated rhythms, whether it be “pushing”, “pulling”, or “in the pocket”. For continuous sounds, we are sensitive to differences within a single waveform. We are able to hear timing discrepancies between two identical sounds playing less than 1ms apart. This effect of comb filtering, whereby sound at multiples of a frequency are accentuated or attenuated, is called “flanging”, coming from the analog era as one put pressure on the flange of an open-reel tape recorder to desynchronize it with another.

We therefore have mechanisms for distinguishing relationships of time where we may lack ability to discern absolute time. Left in isolation, people tend to veer naturally towards a 25 hour day (12). Indicators such as daylight cycles and church bells help ground us in absolute time. Effects like Doppler shift demonstrate that we are sensitive to changes – deltas - in the speed of sound. We know the effect of an ambulance racing by with its siren sounding – the sound itself is a constant melody, but it arrives at our ears at a higher pitch as the vehicle is approaching and shifts to a lower pitch as it drives away. The reason we perceive the sound as such is because the speed of the ambulance

effectively adds to the speed of the sound emanating from it (13). If the ambulance is travelling at 50km/hr, this is about 14m/s. As the ambulance approaches, then, the effective speed of sound of the siren is $340\text{m/s} + 14\text{m/s} = 354\text{m/s}$, while as it departs, the effective speed of sound is $340\text{m/s} - 14\text{m/s} = 326\text{m/s}$. This difference of 8% to the speed of sound is apparent to us as a musically significant pitch shift (14). Time perturbations translating to pitch effects are a natural outcome of one of the fundamental basis of sound: that frequency is inversely proportional to time.

There are times when the combination of the speed of sound, human perceptive accuracy, and performance medium create a situation requiring intervention of non-aural means. The symphony orchestra and its conductor provide one example. The orchestra is a musical ensemble whose goal (more often than not) is to be perceived as playing together. The large number of personnel in an orchestra requires a large stage. This creates a situation where the speed of sound is insufficient to reconcile differences in timing resulting from the stage's dimensions. If a trumpet player in the rear of the orchestra were to follow by ear to play in ensemble with a violinist 10m downstage, the trumpeter would always be late. A distance of 10m at 340m/s creates a delay of 30ms – sufficient to be perceived by the listener in the audience as two separate events. This creates the need for the conductor who directs the orchestra through visual gestures. The musicians no longer depend on the speed of sound for synchronization, but to the speed of light, making a 10m difference of stage seating insignificant. Here is a case where our sensitivity to time in sound surpasses the speed of sound itself.

There are even cases where musical processing surpasses theoretical limits of neural processing time. The speed with which an accomplished pianist can read and perform the score of a rapid passage is not only musically but also physiologically impressive. If one takes the neural transmission time for a piece of visual information to be transmitted from the eyes to the brain, be transformed into neural commands and sent down to the finger muscles, then multiplies this by the number of notes on the page, a rapid passage can surpass the theoretical maximum tempo. There are direct eye-hand coordination as well as semantic phrase analysis processes taking place – the human being dividing nano-time across different parallel processes to achieve an artistic feat of speed.

Our hearing mechanisms use time as a foundational basis. The inner ear is lined with hair cells in the cochlea. Movement of the hair cells trigger nerve firings. The arrangement of these hair cells allows us to distinguish different frequencies (15). This is in fact a temporal mapping – an imprint of time along the topology of the inner ear - that detects different frequencies along the length of the basilar membrane.

Mediums

To this point, we have considered the speed of sound as the speed of acoustic sound waves in air. Musical projects have been realized exploring the transmission of sound in other media such as liquids (16) and solids (17). In the electrical domain, sound pressure waves are transduced to become voltage-based signals, changing the fundamental temporal nature of sound. Audio signals in electrical circuits travel essentially at the

speed of light. Cable length becomes a question of signal strength, and less a problem of timing, for the same reasons given in the orchestra example above.

With digital systems, computation speed becomes the main concern with respect to time. Processing power of early systems was insufficient to treat audio as fast as it came in. The holy grail in that era was to achieve “real time”, and evaluation of a system was based on how many times “out of real time” a process was. As processor power advanced, a one-to-one relationship with nature was no longer the upper limit, and we began to see systems that could boast processing capabilities that were “faster than real time” (19).

In the digital domain, the analog audio signal voltages are encoded in a time-sampled linear binary form. Sound is thus quantized, both in amplitude and in time, 16 bit encoding 44,100 times per second in the case of the Compact Disc. Time in digital form, then, is frozen every 23 μ s. Artifacts of this abstraction of time can be heard in frequency effects. The Nyquist theorem (18) states that the highest frequency that can be encoded is half the sampling rate. Signals exceeding the Nyquist frequency fold back and become audible in an effect called *aliasing*. By fixing an arbitrary minimum time slice, we in effect hear time wrapping around a boundary point.

Combining, or mixing signals from multiple digital sources presents an interesting problem in time – that of synchronizing clocks. Even if time in the two sources is sliced evenly at the same sampling rate, the absolute occurrence of samples in time must coincide precisely for combinatorial data processing to take place. These are all problems inherent in a discrete time representation (20).

Time latency is a characteristic inherent to real time digital signal processing, and can be thought of as the “speed of sound” through computer algorithms. Signal processing inevitably make use of buffers – data caches that are units of temporary storage holding units of data to be processed. The classical depiction of a data buffer is to compare it to a bucket of water with a hole in the bottom. A faucet can feed the bucket, and the hole can empty it at desynchronized rates, as long as the outlet runs fast enough so as the bucket does not spill over. Certain signal processing algorithms require units of source data “looking back at the past” or “looking into the future” to be carried out, creating the need for memory in time. The buffer size defines the base input/output (i/o) latency of the system, and thus the characteristic time delay of the system. It is interesting to note that digital systems that are labelled “real time” are in fact skewed by a constant delay with respect to absolute time.

Creation

These media-specific characteristics described above typically occur to the engineer as technical problematic, or system faults. Artists, on the other hand, can offer a vision of these temporal features not as technical shortcomings, but as characteristics defining the creative potential of a medium. There are tendencies, however, to try to force new time paradigms into old musical models. Part of the worldwide millennium celebrations included a performance of Beethoven’s 9th over the network. The network time latency was dealt with not in a creative way, but in a way consistent with the reallocation of time

described above in the commodity industry. By adding a compensating delay to all parties, this created a situation that would normally be considered musically unsatisfactory.

Artists have instinctively adapted to and exploited physical phenomena and perceptive principles. Composers in the Middle Ages wrote music for the long reverberation times of the church by creating long and slow melodies. They even took advantage of this to sneak in melodies from popular song into their cantus firmus. A tuba player naturally compensates for the slow articulation time of his instrument to play in time with an orchestral *tutti*. J.S. Bach instinctively took advantage of human psychoacoustic stream segregation abilities in writing multivoice counterpoint interleaved into monophonic melody. Drum & bass musicians push rhythm machine programming to our perceptual limits of discerning discrete events with their ultra-rapid snare drum fills. While Cage's 4'33", in its absence of sound and clear definition of time, heightens sensitivity to our aural surroundings, the full frequency nature of Japanese Noise music works to suspend our sense of time. Farmersmanual brings parallel processing dynamic onstage to create non-eventlike states in concert (21). Transmission delays will be considered a hindrance as long as we try to superpose old musical forms onto the network. Rather, a new musical language can be conceived, respecting notable qualities of the medium, such as packetized transmission and geography independent topology (22, 23). These features can be said to define the "acoustic" of the network, a sonic space that challenges existing musical notions of event, authorship, and time.

Whether working in relation to or in the absence of structure, physical or formal, music and sound art take on architectural qualities. Composers create temporal forms often inspired by physical structures (24, 25, 26). Improvisers depend on structures of time to free themselves from form. Artists place or displace sound in social and physical contexts. The human ear is a multichamber mechanism of sound detection, amplification, and sensation. The coordinate system characterizing sound is one of space and time – domains that at base, we cannot touch. Music, Sound art, and aural perception are architectures along these axes, in realms that are intangible, that nonetheless shape and define a tangible universe. The speed of sound is the fundamental element, for velocity is a function of distance and time.

Atau Tanaka
Paris, Tokyo
September, 2000.

References

1. Boulez, P. 1990. *Orientalisms: Collected Writings*. Cambridge: Harvard University Press. <http://www.hup.harvard.edu>
2. Russolo, L. 1913. "L'Arte di rumori." transl. 1986. *Monographs in Musicology No. 6*. Hillsdale: Pendragon Press. <http://www.pendragonpress.com>
3. Cage, J. 1961. *Silence*. Middletown: Wesleyan University Press. <http://www.wesleyan.edu/wespress/>

4. Labelle, B. 2000. "Sounding Out: Reverberations Across Social Space." *Just About Now*. Rotterdam: V2_archief. <http://www.v2.nl/archief>
5. EnviroMeasure. 1997. *Speed of Sound Calculator*.
http://www.measure.demon.co.uk/Acoustics_Software/speed.html
6. Cramer, O. 1993. "The variation of the specific heat ratio and the speed of sound in air with temperature, pressure, humidity, and CO2 concentration." *J. Acoust. Soc. Am.* 93(5): 2510-2616. <http://asa.aip.org/jasa.html>
7. Virilio, P. 1977. *Vitesse et Politique*. transl. 1986. *Speed & Politics*. New York : Autonomedia. <http://www.autonomeia.org>
8. Schafer, R. M. 1977. *The Tuning of the World*. New York: Knopf.
<http://www.randomhouse.com/knopf/>
9. Helmholtz, H. 1863. *On the Sensations of Tone*. New York: Dover.
10. Blauert, J. 1996. *Spatial Hearing*. Cambridge: MIT Press. <http://mitpress.mit.edu>
11. Winckel, F. 1967. *Music, Sound and Sensation*. New York: Dover.
12. Coren, S. 1997. *Sleep Thieves*. New York: The Free Press.
<http://www.thefreepress.com>
13. Pierce, J. R. 1992. *The Science of Musical Sound*. New York: W.H. Freeman & Co. <http://www.whfreeman.com>
14. Partch, H. 1979. *Genesis of a Music*. Cambridge: Da Capo Press.
<http://www.dacapopress.com>
15. Roederer, J.G. 1995. *The Physics & Psychophysics of Music*. London: Springer-Verlag. <http://www.springer.de/>
16. Redolfi, M. 1996. *Detours*. Nice: Mirage Musical.
17. Tsunoda, T. 2000. "Monitor Units for Solid Vibration." *Sound Art – Sound as Media*. Tokyo: NTT InterCommunicationsCenter. <http://www.ntticc.or.jp>
18. Nyquist H. 1928. "Thermal Agitation of Electric Charge in Conductors." *Phys. Rev.* 32:110-113.
19. Puckette, M. 1991. "FTS: A Real-time Monitor for Multiprocessor Music Synthesis." *Computer Music Journal*. 15(3):58-67. <http://mitpress.mit.edu>
20. Pohlmann, K. *Principles of Digital Audio*. New York: McGraw-Hill.
<http://www.mcgraw-hill.com>
21. Farmersmanual. 2000. *off-ICMC*. Berlin: Podewil. <http://www.podewil.de>
22. Tanaka, A. 1999. "Netmusic - a Perspective." *Festival du Web*. Paris : Webart.
<http://www.webbar.fr>
23. Tanaka, A. 1999. "Network Audio Performance and Installation." *Proc. Intl Computer Music Conf.* San Francisco: ICMA. <http://www.computermusic.org>
24. Treib, M. 1996. *Space Calculated in Seconds*. Princeton: Princeton Univ. Press.
<http://pup.princeton.edu>
25. Xenakis, I. 1971. *Formalized Music*. Hillsdale: Pendragon Press.
<http://www.pendragonpress.com>
26. Lowman, E. L. 1971 "Some Striking Proportions in the Music of Bela Bartok." *Fibonacci Quarterly* 9(5): 527-528.
<http://www.sdstate.edu/~wcsc/http/fibhome.html>