[**This is the second draft of this Appendix begun on 25 August 2008**]

Appendix R: Human Biosonar Capability

James T. Fulton
jtfulton@cox.net

This Appendix is a work in progress. It is designed to evaluate the potential for human echolocation by comparing the human physiology and linguistics with that of the highly optimized system of the Bottlenose Dolphin and other members of the Cetacea family of whales and dolphins. It is also designed to suggest some scenarios that might be useful in training humans, particularly those with vision impediments, to use echolocation more effectively.

R.1 Background

It was recently brought to my attention, through the television program focused on a totally blind young man who was able to ride a bicycle and traverse familiar territory with extraordinary skill based on the use of his acute hearing and augmented by vocalizations designed to provide echolocation information.

I was simultaneously working on the echolocation and possible linguistic skills of the Bottlenose Dolphin. It appears it is possible to develop the skills illustrated by Ben Underwood even more fully. This would involve starting with blind children at the earliest practical age. This in turn would involve developing a training program designed to familiarize them with the potential of echolocation as soon as possible so that the plasticity of their brain can be employed to expand their intrinsic capability.

R.1.1 Sources

This discussion will draw heavily on a sister Appendix L, "Dolphin Biosonar Echolocation: A Case Study" available at www.hearingresearch.net/pdf/Dolphin biosonar echolocation.pdf . The amount of data available from dolphins is limited and the data must frequently be interpreted in order to utilize it. Dolphins are very intelligent animals. When investigators design straightforward experiments to obtain pertinent information, the dolphins invariably respond by adopting more sophisticated strategies and sound sequences that the investigator expected.

Wikipedia provides a wealth of information on sounds associated with the human voice. The material on clicks is particularly useful, http://en.wikipedia.org/wiki/Click_consonant. A link at the end of that page is also very useful, Collection of click-language links and audio samples . Excerpts from this material are provided in Section R.9 of this document.

Blesser & Salter provided a useful reference relating to the potential of environment perception using human auditory facilities[1]. Along with many useful and specific definitions in the introduction, it explores many aspects of auditory spatial awareness in Chapter 2. They note "the ability ot create an internal picture of external objects and geometry is greatly enhanced when strong motivation, greater than average skills, and extended opportunity to practice are present (page 40)." Their discussion of echoes versus background sound as scene coloration may need more discussion, especially their quoting a minimum 10 meter range criteria for an echo on page 42 but noting the ability of blind individuals to identify shapes at distances as small as two feet (presumably using active echolocation) on page 44. They suggest a typical aural space is characterized by its aural embellishments (page 52). They also engage in a valuable discussion of the analogy between visual and aural reflections (page 55).

After noting that echolocation has come to mean both passive and active spatial awareness in the vernacular, Blesser & Salter focus primarily on aural "imaging" of the nearby environment based on ambient background sound rather than active echolocation. They focus on separating the acoustically reflective characteristics of materials from the intrinsic sound falling on such materials, much like the reflective optical characteristics of material illuminated by a light source. Thus, their work pertains primarily to materials and shapes whether actively or passively illuminated acoustically. They associate aural perceptions as involving coloration if due to distributed quasi-continuous acoustic sources and as echos if due to impulse-like sources. They discuss cognitive maps based on acoustic reflections in the same context as those based on light reflections. They also develop the subtle differences between visual and acoustic reflections from the experiential perspective (pg 55). In their broad use of the term echolocation, they note, "sensory skills are acquired, rather than innate; they are based on personal utility and lifestyle."

---

[1] Blesser, B. & Salter, L-R. (2007) Spaces Speak, Are You Listening? Cambridge, MA: MIT Press

This analysis is focused primarily on active echolocation but can be extended or merged with other passive target location techniques.

R.1.2 Potential number of blind able to use echolocation

Specifically who are we trying to help out of the total pool of “blind people,” Kish has provided some estimates of the blind population in the USA. I have re-tabulated his estimates and added additional numerics. .

1% of total population –those considered legally blind. 3 million of 300 million.

60% of legally blind are over the working age of 65.

5-10% of legally blind are under the working age. 150,000-300,000

~30% of legally blind of working age.

10-20% of legally blind are functionally blind. 300,000-600,000

4% of legally blind have been significantly blind since birth. 120,000

10% of the above 4% being functionally blind since birth. 12,000

60% of those significantly blind since birth have other neurological problems. 72,000

40% of those significantly blind since birth can be considered “vanilla blind,” no other major neurological problems. 48,000

These numbers suggest there are about 1,000 babies a year born functionally blind and without other significant neurological limitations. They are presumably alert, mobile and potentially motivated. Are these the subjects we are trying to help? Is the goal to help those in a larger pool as well?

It is reasonable to assume those with complex neurological problems centered on the sensory modalities suffer from disease of the diencephalon of the midbrain rather than the cerebral cortex.

R.1.3 Glossary

Blesser & Salter have provided a number of very clear definitions related to this analysis. They begin with a useful delineation between the terms acoustic and aural.

Acoustic– the behavior of sound waves (vibrations) in solids, liquids and gases.

Aural– the experience of a sonic process.

Hearing– the detection of sound.

Listening– active attention or reaction to the meaning, emotions and symbolism contained within sound.

They used the above definitions to develop the following additional terminology.

Aural architecture– the properties of a space that can be experienced by listening.

Aural embellishment– An acoustic object or geometry that produces aesthetically recognizable acoustic attributes that add aural richness and texture to a space.

Acoustic architecture– A technologist who designs spaces to accommodate the requirements of aural architecture.

Spatial acoustics– the way that a space changes the physical prpoerties of sound waves.

Cultural acoustics– the way that listeners experience sound within a space.

R.2 Overview of echolocation in the Bottlenose Dolphin

R.2.1 Geometry of dolphin echolocation

R.2.2 Waveforms of dolphin echolocation

Lemerande has provided an excellent set of data for the bottlenose dolphin[2]. **Figure R.2.2-1** provides a recording of a single click from this dolphin. It is nearly ideal, showing one complete sine wave of high intensity followed by an oscillation consisting of two smaller positive going peaks.

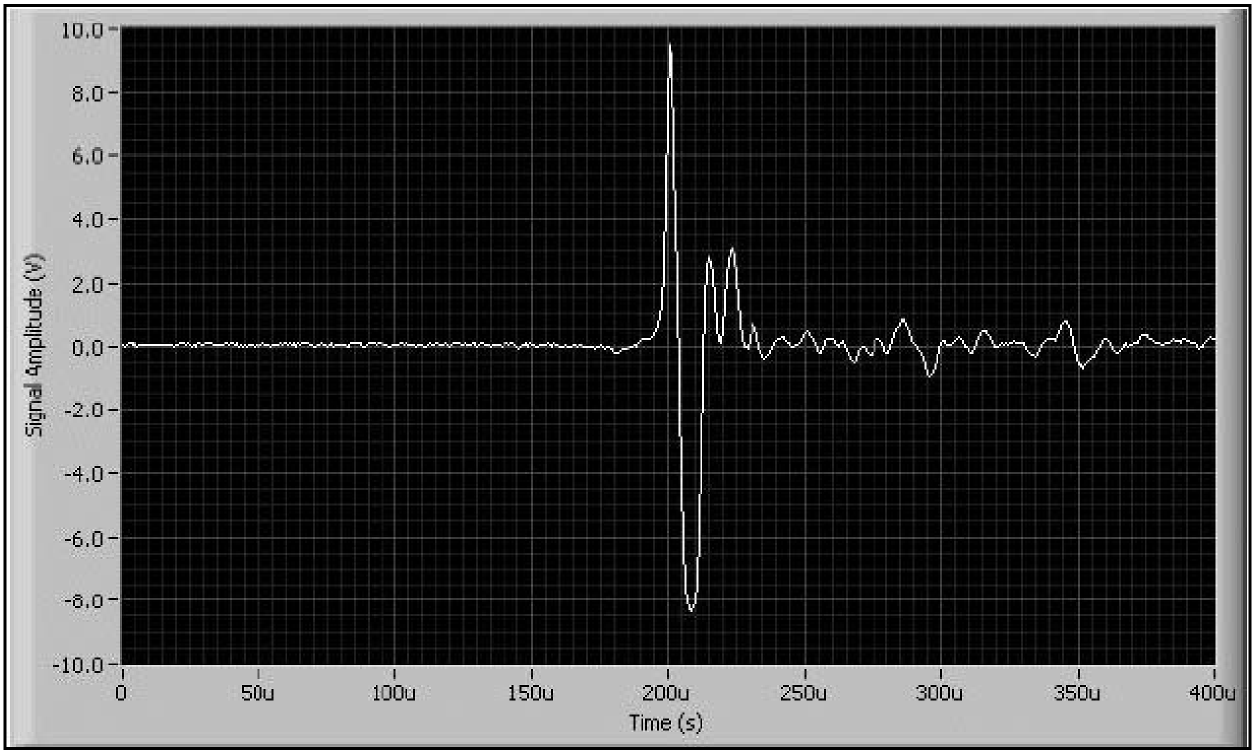

Figure R.2.2-1 Nearly ideal click waveform from the bottlenose dolphin. From Lemerande, 2002

The click waveform displayed above is usually found in a train of pulses as displayed in **Figure R.2.2-2**. The click

---

[2] Lemerande, T. (2002) Transmitting beam patterns of the Atlantic Bottlenose Dolphin (*tursiops truncatus*): Investigations in the existence and use of High frequency components found in Echolocation signals. Naval Postgraduate School, Monterey, CA 93943-5000

train consists of an equally spaced set of clicks of the type shown above. The frequency of these click trains can vary from 75 to 250 clicks per second. This suggests the clicks are generated by a mechanism not greatly different from the vocal chords of humans which have a similar fundamental frequency range. The highest amplitude component is the initial click. It is followed by the lower amplitude trailing waveform. Note in this sequence, the animal reduces the amplitude of its clicks as it converges on the target on the right. Some variation in signal intensity is present as the animal nutates its head back and forth during the closing pursuit. As noted elsewhere, it is also possible the animal is able to steer its forward pointing high frequency acoustic beam over small angles by changing the shape of the melon in its forehead (which acts as an acoustic lens).

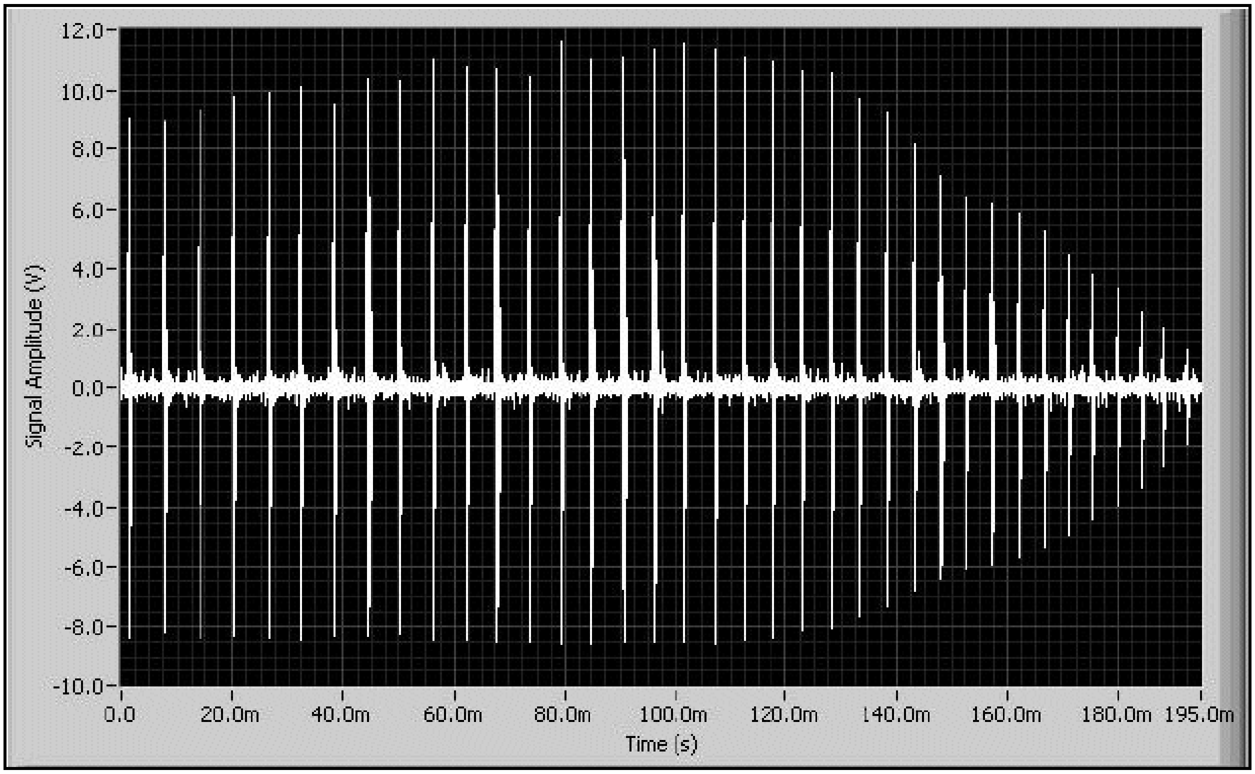

Figure R.2.2-2 Click train of bottlenose dolphin. From Lemerande, 2002.

R.2.3 Resolution capability of the dolphin

Extensive experimentation has been undertaken to determine the range resolution of the bottlenose dolphin. It appears the animal can differentiate reliably between the presence or absence of adjacent aluminum disks on the order of a few centimeters in diameter at 100 meters in water. It appears they have a similar range resolution but this value is controversial. Au has suggested the range resolution for static targets is on the order of 1.0-1.1 cm. Similarly, it appears from calculations that they can employ the Doppler Principle to determine changes in range velocity on the order of 1.5 miles/hour. Their ability to track targets at lesser ranges is known to extend down to distances only slightly beyond the end of their rostrum (mouth).

R.2.4 XXX

R.3 XXX

R.4 A comparison of the dolphin and human echolocation capabilities

Figure R.4.1-1 provides a simple comparison of the echolocation potentials of the Human and the Bottlenose Dolphin. There are two major differences that are difficult to illustrate. The velocity of sound in water is about 5 times higher than in air. The maximum usable frequency the ears can sense for echolocation in dolphins is believed to be about 150 kHz whereas in humans it is much lower, probably on the order of 15 kHz. The vocal capability of the dolphin is compatible with the 150 kHz value while that of the human is concentrated at a lower frequency, probably in the 3-5 kHz region. The forward pointing high frequency beam of the dolphin has been measured at about xxx degrees full angle. The equivalent beam for the human is highly dependent on the shape of the mouth and lips but is always broader than that of the dolphin at high frequencies. The larger aperture and the front facing direction of the dolphins ears provide it with a much higher sensitivity based on geometry alone. The spacing of the two ears appears to be similar in the two species.

The bottlenose dolphin generates a high frequency beam that is ten degrees wide at the 3 dB points and 120 kHz. Each ear has a receiving pattern that is nominally ±26 degrees at its first null. The ears are independently pointable by changing the contour of their outer surface. However, when presumably pointed forward, they have been measured to have a divergence of 26 degrees between their axes.

The vocal capabilities of the two species are similar at the simple box level of illustration but their implementations are quite different. Both employ a larynx to produce tones and very complex variants of tones and tone combinations. In the human, this capability is augmented by the great flexibility provided by the flexible oral cavity and its elements, including the highly maneuverable tongue. The dolphin may use the nasal cavity in a manner similar to the human oral cavity but it appears less flexible at low frequencies based on spectrographic data. The dolphin employs a set of structures in its nasal cavity to generate high frequency sounds similar in concept to the way the human can generate staccato sounds (clicks) using its tongue and palette. However, the dolphin does it more efficiently and at greater effective sound intensities.

It is worth noting that the dolphin does not share a common breathing and feeding path like the human does. No sound issues from the dolphins mouth. It is either radiated out through the melon of its forehead at high frequencies or is conducted in the airstream issuing from its blowhole at low frequencies. The dolphins at the aquarium are trained to open their mouths when communicating with humans vocally because that is what the human expects to see.

The connection between the outer ears of the dolphin and human are quite different. The energy from the tympanic membrane at the end of the outer ear canal connects to the vestibule of the labyrinth via the oval window (and stapes). In the dolphin, the energy from the outer ear is applied to a different surface of the vestibule of the labyrinth. The oval window (and stapes) have atrophied after the animal returned to the sea. The apertures on the side of the dolphins head are no longer used as ears but may be used as a depth gauge or velocity measuring device under water.

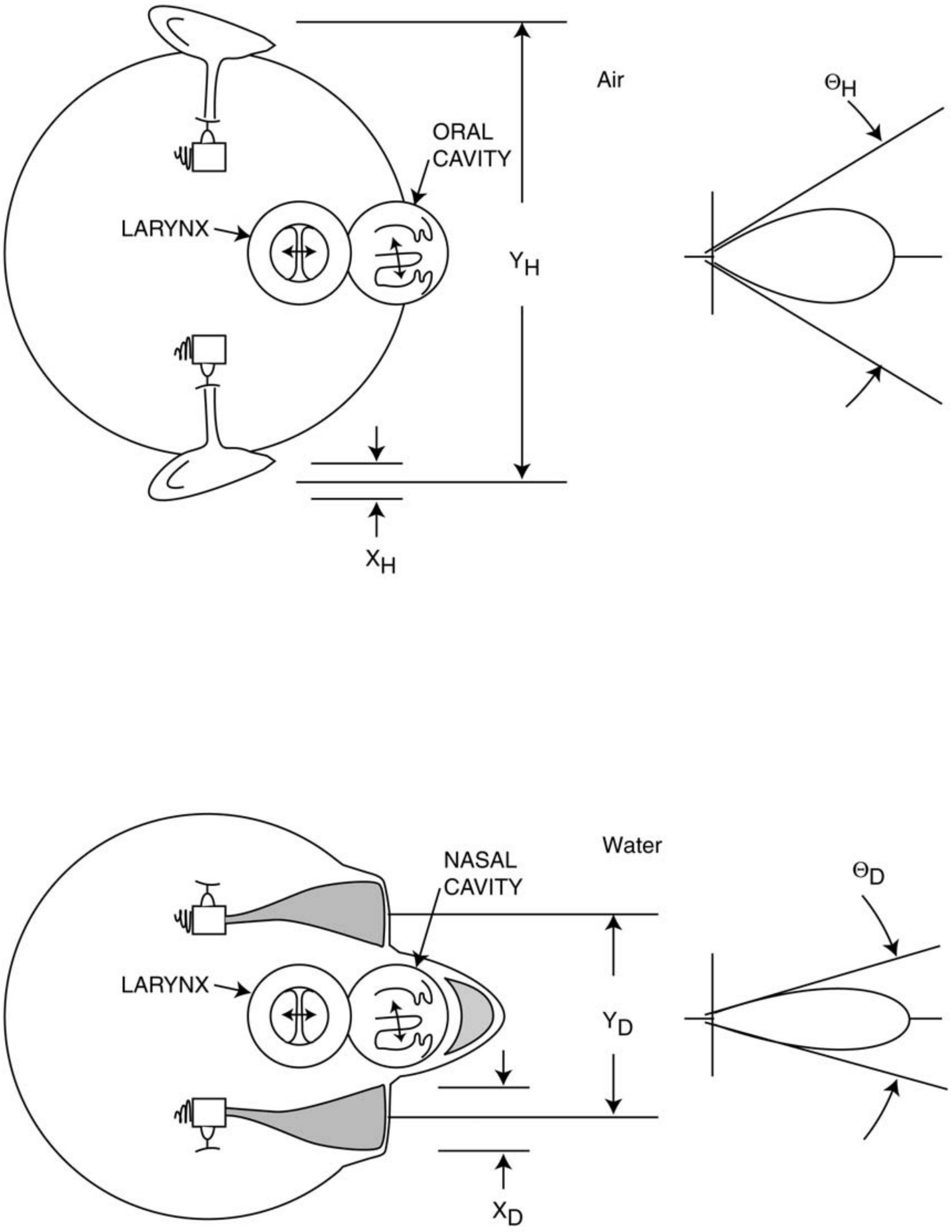

Figure R.4.1-1 A comparison of the human and dolphin echolocation apparatus. The width between the ears is similar, YD ≈ XD ≈ 8 inches. However, the forward projection of the dolphin ear is about 3 inches in diameter while the front projection of the human ear is on the order of 1 inch in diameter. The total angular width of the high frequency beam formed by the dolphin is about 10 degrees at the 3 dB point (& 120 kHz) while that of the human is broader and variable.

R.4.1 Reported passive discrimination capabilities of humans

Blesser & Salter have given a series of values for the reporting of object size versus distance (page 44) in what was apparently natural environments without any active sound from the subject.

> “Rice (1967) showed that listeners can detect a difference of 1 centimeter in a 9-centimeter disk at a distance of 60 centimeters (2 feet). Kellogg (1962) showed an even higher level of discrimination: listeners detected an area difference of 5 square centimeters on a square of 60 square centimeters at a distance of 2 meters (7 feet). One blind subject could reliably detect a 1-inch disk located at a distance of three feet (Rice, 1969-70). Even more remarkably, Hausfeld et al. (1982) demonstrated that listeners could distinguish square, circular, and triangular objects. One blind subject was able to recognize a stop sign by its hexagonal shape. Kellogg (1962) found that on the most difficult discriminations tasks blind individuals performed significantly better than sighted subjects who were blindfolded.”

R.5 Operation of the dolphin echolocation system

It is generally believed among the dolphin research community that the dolphin, and particularly the bottlenose dolphin, “sees” an image of the field illuminated by its high frequency vocal system. The size of the field illuminated by this dolphin vocal system is quite similar in angular extent to that of the foveal portion of the human visual system (the central 1.2 degrees of the visual field). It is also sensed by a binaural hearing system in analogy with the stereographic capability of the visual system of both the human and the dolphin.

The dolphin is able to determine the velocity of the objects in its high frequency field of view with a minimum value on the order of two kilometers/hr (1.5 miles/hour).

What the dolphin actually sees in its high frequency auditory world has not been demonstrated. However, for a static scene, the data suggests it sees a nominally black and white 3-D image of a field of about xxx elements in both horizontal and vertical dimensions. Axial motion within that scene above the minimum threshold are seen as either positive or negative values that may be analogous to color in the human visual system. Under this analogy, a school of fish swirling about a vertical axis would be images as a three-dimensional cluster with the side swirling toward the dolphin appearing blue and the side swirling away from the dolphin appearing red.

The information sensed by the high frequency auditory system of the dolphin is presented to the auditory perigeniculate nucleus (a portion of the diencephalon or midbrain) where it is assembled into an image using a multidimensional Cartesian array of neurons. This acoustic array is similar to that used in the visual system of both dolphins and many other mammals. In the human, the visual system delivers information from the fovea to the visual perigeniculate nucleus where it is processed in a similar multidimensional Cartesian array of neurons. The visual and auditory perigeniculate nuclei are adjacent to each other within the diencephalon. It is possible, but unverified, that these two perigeniculate nuclei share a common substrate or multidimensional array. The auditory perigeniculate nucleus is adjacent to the inferior colliculus. In the literature, these two elements are not differentiated and it is common to find references to the very large inferior colliculus in the dolphin family.

The reader will note, the information received from the fovea of the eyes is not delivered to the so-called “primary visual cortex.” While this title may refer to the first or initial portion of the visual cortex, V1, it does not properly describe the visual processing area of primary importance to the human, the visual perigeniculate nucleus.

The similarity between the auditory perigeniculate nucleus and the visual perigeniculate nucleus in the dolphin, and the similarity between these two components in the human brain suggest strongly two ideas.

First, that the dolphin does “see” in the high frequency portion of its auditory system.

Second, the human has a similar but less developed intrinsic capability. The auditory perigeniculate nucleus is much larger and more highly developed in the dolphin than is the equivalent structure in the human. But, this is not to say the human auditory perigeniculate nucleus cannot be developed more fully through use.

Furthermore, it is possible that through the plasticity of the brain a portion of the unused visual perigeniculate nucleus could be requisitioned by the auditory perigeniculate nucleus to enhance the overall capability of the child.

> **It is the goal of this study to explore how the unused portion of the visual perigeniculate nucleus can be requisitioned by, and used efficiently by, the auditory perigeniculate nucleus in order to achieve a more capable object location capability in the blind child.**

R.6 Operation of the human source location system

The human sound-based source location system consists of a passive and an active component. The passive capability will not be addressed in this document. The active component can rely upon an internal sound source, in which it can be described as a monostatic system (source and receiver in the same enclosure), or an external sound source, in which it can be described as a bistatic system (source physically separate from the receiver). In the bistatic case, there are two options. The source can generate a signal at a random time unknown to the receiver in advance, or it can generate a signal under the control of the overall system. A system generating a signal at a random time will be called asynchronous. A system generating a signal at a known time will be called synchronous.

Dan Kish has defined what he calls a SoundFlash device. To the extent, the signal generated by this device is controlled by the user, it will be considered a synchronous bistatic device. If it is controlled by an investigator independent of the receiver, it will be considered an asynchronous bistatic device. A sound created by the user using his vocal apparatus will be considered as coming from a monostatic system.

> This investigator considers it preferable if the user employs his own vocal apparatus as a source, a monostatic source, where possible. For young children this may not be possible initially. In that case, it is considered preferable to use a bistatic source producing a signal comparable to that eventually produced by the child.

It is probably obvious that the brain can easily calculate the time difference between a vocally generated sound and its received reflection. This calculation is the primary calculation in monostatic echolocation. While it is quite possible for the subject to perform a similar calculation based on a reflected signal received after a sample of the original signal is received in a bistatic situation, it involves a more complex operational situation and a generally less precise calculation.

{*The brain uses a variety of processing algorithms to determine the position of objects in its environment. Only the relatively simple algorithms will be discussed here*.)

The brain uses the time delay between the initial signal generated and the received echo(s) from that signal to compute the distance to the object(s). By comparing the time difference between the signals from the two ears, it is able to determine the angular position of the object(s). The system will rely on the leading edge of the generated and reflected signals if they are of adequate intensity. If not, it will attempt to integrate the energy in the source and reflected signals in order to achieve the same result. Ideally, the generated signal would be a single pulse consisting of a single sinusoid waveform at an optimum (generally high) frequency. Such a source signal allows the simplest calculations relating to the location of the reflection(s) and some additional information relating to the acoustic characteristics of the reflecting material(s). Frequently, it is necessary to transmit a signal consisting of more than one sinusoid (a frequency burst) to obtain adequate power in the signal to achieve the desired maximum range. It is still preferred to generate a signal at only one frequency if possible in order to achieve maximum range and range rate (velocity) information concerning the reflected signal(s).

While a variety of information is available on the ability of the human vocal system to create clicks in the context of speech and communications (Section R.9), the applicability of this data to echolocation is limited. It is proposed that the ideal signal for human echolocation is an alveolar click generated where the tongue is quickly withdrawn quickly from the front of the palette while the vocal cavity is blocked at the throat. This process can form an area of reduced pressure within the oral cavity. This rarefaction of the air pressure is followed by a raising of the air pressure (a compression). The result is a single sinusoid waveform followed by a series of artifacts. The shape of the mouth and lips are critical in determining the shape and the maximum intensity of the final waveform projected forward of the mouth. This is the area where more experiment is needed to optimize the choice of alveolar click.

R.6.1 The capabilities of Ben Underwood

A recent documentary highlighted the capabilities of Ben Underwood, a person in his high teens who has been totally blind since the surgical removal of his eyes at an early age. Ben has developed a remarkable capability involving both passive and active echolocation using both asynchronous bistatic and his own synchronous monostatic sources.

Ben has developed his asynchronous bistatic capability to an exceptional degree, as have other sightless people. He perceives walls, and other terrain features quite effectively, based on the presence of other bistatic sources or even his own incidental sound sources (such as the sound from his shoes). He is also able to perceive some elements, such as doorways, based on essentially the ambient background noise reflecting off of the surface surrounding the doorway. It is inspiring to see Ben riding a bicycle by himself in a protected, but not empty, cul-de-sac near his home.

Watching and listening to Ben Underwood demonstrate his synchronous monostatic capability briefly in the TV documentary, it is difficult to determine whether he is using an alveolar click or a dental click (formed by withdrawing the tongue quickly from the back of the front teeth). It is also difficult to describe the shape of his mouth in detail based on the short video. Ben was obviously achieving a degree of echolocation at ranges as short as 10 inches (25 cm). He appeared to be able to tell the difference in crude shapes at that range after a moment of concentration.

These capabilities were developed by Ben essentially on his own (with great support from his acoustically untrained Mother) beginning around age five.

> **It is proposed, and appears obvious, that he, and similar subjects, could have developed an even higher capability if he had been provided more technical guidance during his earlier years**.

R.6.2 Background on the plasticity of the brain

An excellent overview of the plasticity of the brain was assembled for the lay press by Sharon Begley of the Wall Street Journal on Dec 2, 2005. It is reproduced in full at the end of this appendix.

Steven & Blakemore have provided a scientific paper providing more detail than Begley[3]. Their citations lead to the conclusion that the human brain remains more plastic than previously believed.

Thaler, Arnott & Goodale have recently provided an extensive paper on the fMRI results of experiments in echolocation involving an early blind participant (Daniel Kish), a late blind participant (Brian Bushway) and two normal sighted individuals[4]. The results clearly show that Daniel has taken advantage of the plasticity of his brain at an early age and that he is fully utilizing the right occipital lobe (area 17, or V1) to process acoustic signals. Brian was blinded at age 14 and does not show the same plasticity or utilization of either lobe of area 17.

R.7 Potential programs to improve blind children's capabilities

The human auditory system, in conjunction with its vocal system, is demonstrably capable of achieving a useful degree of active echolocation. It also appears the auditory system is capable of a poorly documented capability to perform passive object location, based on subtle changes in the ambient sound levels encountered near structures. It appears that these capabilities can be enhanced through training, and particularly training at an early age when the unique learning capabilities of the human appear to be available.

This investigator is not qualified to design the appropriate training protocols. However, he might be able to offer some suggestions that could be included in a properly designed program.

It has been suggested that a program similar to the "Lilli Nielsen boxes," crib-sized enclosures that bring a large variety of sensible objects to within reach of the infant while flat on its back.

Useful sites discussing these boxes include:

---

[3] Steven, M. & Blakemore, C. (2004) Cortical plasticity in the adult human brain *In* Gazzaniga, M. *ed*. (2004) The Cognitive Neurosciences, 3rd Ed. Cambridge, MA: MIT Press Chap. 89

[4] Thaler, L. Arnott, S. & Goodale, M. (2011) Neural correlates of Natural human echolocation in early and late blind echolocation experts *PloS ONE* vol 6(5), pp e20162 (16 pages)

http://en.wikipedia.org/wiki/Lilli_Nielsen

www.lilliworks.com/aln-dec01.htm

www.wonderbaby.org/articles/play-area.html

The following subsections will describe potential elements of several training protocols. It must be left to others to determine if these elements are useful. It must also be left to others to determine what children have sufficient capability to take advantage of this potential training.

R.7.1 Optimum age for beginning ehcolocation activities

This section needs development by someone more skilled in child development. But the comments in Carter's book on "Mapping the Mind" offers several pertinent thoughts[5]. She notes the following. "'Proper' language starts in the second year with the activation of the two major speech areas that occupy separate but neighboring areas on the side of the brain. One – Wernicke's area – is specialized for language comprehension, while the other – Broca's area – deals with speech articulation." Both of these areas are under the control of the thalamus within the diencephalon. "One of the clearest precursors to language is babbling – the speech-like torrent of sounds that babies typically start generating at about eighteen months. This is followed within a few months by a rapidly expanding vocabulary of proper words." Is this the key time to begin serious efforts to develop echolocation skills? Carter goes on to discuss a seriously deprived teenage child due to near absolute isolation. The child had not learned a native language. When she finally encountered spoken words, fMRI scanning showed "her brain processed them *in an area normally reserved for environmental noises* (emphasis added)." Carter also notes, citing Kim[6], that second languages are typically processed in a different language area of the brain than is the native tongue. While this may not be true for second languages acquired simultaneously with the first language, it is indicative of the ability of the child to acquire additional skills, probably including echolocation, without restricting its ability to learn a primary language.

Carter also noted (pages 173 & 177-180), the condition known as synaesthesia, wherein a person perceives a taste sensation while hearing a particular sound, or a musical sensation while viewing a color. These conditions suggest a certain sharing of sensual inputs within the brain in unusual ways. It is possible that these can contribute to a blind child imaging a set of auditory reflections much like a dolphin is believed to image high frequency auditory reflections.

If Carter's observations are correct across the population, they suggest a blind child can be most effectively helped to learn echolocation skills beginning at about two years of age – generally beyond the stage where a simple Lilli Nielson type box, or little room, would be effective. The child achieves a considerable degree of mobility at this same age, and their degree of inquisitiveness about their environment also mushrooms. Thus, a more expanded concept of a specialized enclosure focusing on passive and active source location, and compatible with a two-year old child might be optimum.

R.7.2 Scenario #1 – A crib scenario

For a child with very limited or no sight, the task calls for a distinctly separate "little room" from that designed for a child with nominal sight. This is not to say the conventional room is not also valuable in such a situation for other purposes.

The goal of the hearing augmentation room should be to focus exclusively on augmenting the child's hearing in the context of passive object location and active echolocation. This calls for an uncluttered room with objects specifically designed to associate the objects present with their auditory signature, whether active or passive.

[The following sentences assume, the alveolar click is shown to be the most useful click for echolocation in terms of power output and range discrimination capability.]

As soon as a child develops the ability to perform an alveolar click, an effort should be made to use this click, either

---

[5] Carter, R. (1998, paperback 2000) Mapping the Mind. London: A Phoenix paperback

[6] Kim, K. et al. (1997) Distinct cortical areas associated with native and second languages. *Nature* vol 388:6538, pg 171

recorded or precisely simulated, as the nominal sound source associated with the hearing augmentation room.

One strategy should be to have a toy, in an otherwise empty hearing augmentation room, slowly approaching the child in a linear fashion while generating the nominal alveolar click.

A second strategy should be to have the toy operate as described above, while a separate sound reflecting surface is located in a fixed position within the hearing augmentation room.

A third strategy is to have the separate sound reflecting surface located in a fixed position within the hearing augmentation room while the child is making alveolar clicks. Later two or multiple reflecting surfaces can be introduced to familiarize the child with multiple reflections from a single source (whether from an external source, the above toy, a separate external source located above the head but in a similar geometric position as the mouth, or eventually its own alveolar clicks.)

R.7.3 Scenario #2 – A small child's room

If the developmental psychologists concur with the earlier quotations from Carter, it is useful to consider a room optimized for a visually impaired child of at least 15 months of age. Such a room would be optimized to provide a wide range of acoustic features that the child would be immersed in through out its day. In concept, the room would consist of:

1. A hard floor without covering (except in selected area to introduce other levels of sound reflection).
2. Hard walls, possibly using a Formica or hard lacquered wood wainscot extending at least three feet from the floor.
3. A doorway that when open exposes a hall with much less (but not zero) acoustic reflectivity than the wainscot.
4. Panels a few feet wide or less covering portions of the wainscot with lower reflectivity materials.
5. Various free standing and movable panels with high acoustic reflectivity (possibly incorporated into the furniture).
6. A small-waterfall or burbling brook device located so as to generate sound as well as reflections of that sound from portions of the wall or free-standing panels. The device to be used intermittently, but frequently and for extended periods after the child becomes acquainted with it. Its loudness and position chosen so that it does not dominate more than a small portion of the room.
7. A motorized clacking device, such as the xxx, that produces a continuous intermittent sound in one corner of the room, and is reflected off of at least one other hard surface. Used intermittently, but frequently for extended periods after the child becomes acquainted with it. Its loudness and position chosen so that it does not dominate more than a small portion of the room. Ideally, this clacking device would simulate the alveolar click that the child can make using its tongue and front palette.
8. The child's shoes should have a hard bottom surface, possibly even taps, so that it is continually exposed to the echos associated with its footsteps immersed in the reflective acoustic environment.
9. A push or pull toy that makes a cluck type sound when it moves (not the new toys that make realistic barking or verbalized sounds). A pull-dog called Pluto was available when I was young. There appear to be a number of pull-ducks available now that provide a simple quack.

Such a room can familiarize the child with the world of acoustic signatures that it is learning to live with (in addition to mere voices and recorded programing generally devoid of reflection components due to the movement of the child or the source). It can also acquaint the child with the "clacking sound" that he/she will soon learn to generate vocally and begin to use as an effective echolocation tool as it expands its world beyond its room.

The parents might want to consider introducing a clucking sound into their repertoire when entering or leaving the room so the child becomes familiar with the dynamics of the room related to those sounds.

R.8 XXX

R.9 Linguistic characteristics of a click

R.9.1 The anatomy of human clicks

Ladefoged & Maddieson have provided **Figure R.9.1-1** showing the names given to the articulators and articulator targets in human sound production[7]. The heavy lines on the left show regions of the tongue that normally move together even though one part might be considered the leading or dominant part. This feature includes the paired lips. The odd numbers refer to tongue movements dominated by the tip of the tongue. Ladefoged & Maddieson provide a more detailed figure differentiating between these odd numbered movement and the even numbered movements dominated by the blade of the tongue. The lines marked 7, 9 & 11 on the right show the regions of tongue motion of most interest in studying clicks. Although the position of moveable structures away from the midline can be important in differentiating between sound, the most significant articulary characteristics are associated with the midline.

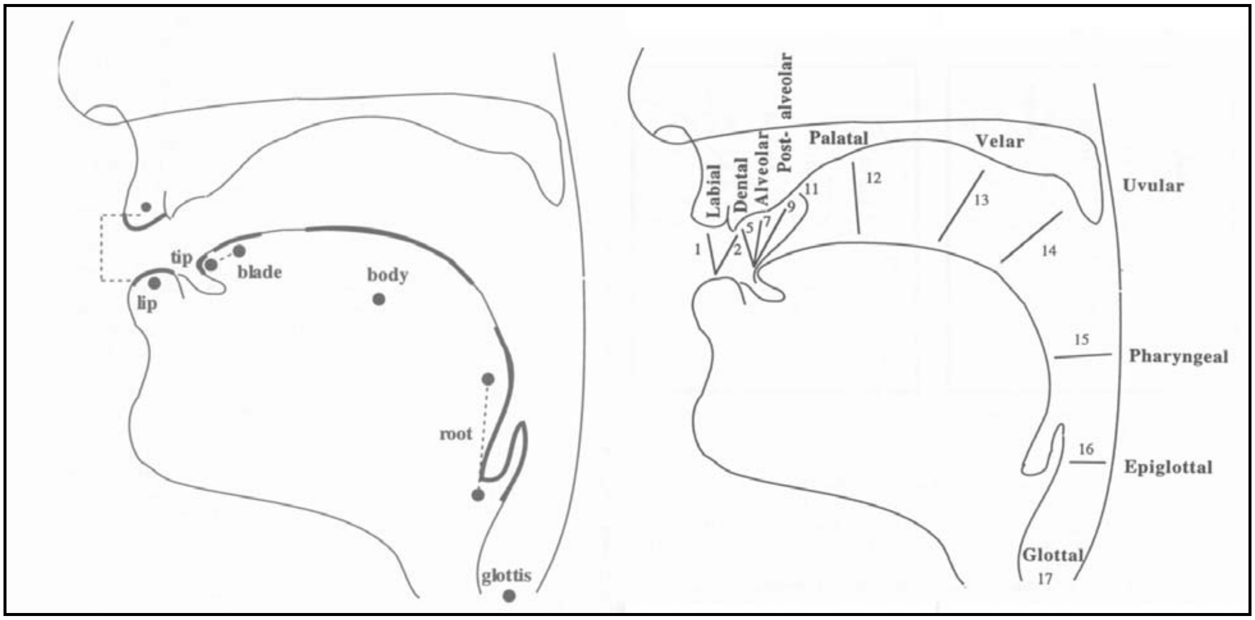

Figure R.9.1-1 The five movable structures forming active articulators and their nine targets of the vocal tract. The lips are treated as a pair on the left. From Ladefoged & Maddieson, 1996.

Figure R.9.1-2, also from Ladefoged & Maddieson, shows the formation of the two clicks of most interest in echolocation. In both the alveolar and palatal clicks, the tongue must close off the airway at two points along the midline, in the alveolar or palatal area and in the velar or uvular area. The click is formed by reducing the pressure on the air trapped between these closures and then retracting the tongue to allow air to rush into the low pressure cavity. In these diagrams, the greatest change in air volume associated with the click is seen in the alveolar click. It is presumed that this click has the potential to be the loudest.

As an aside, it should be noted that the hearing system is most sensitive to, and reacts earliest to, a sudden lowering of air pressure within the ear canal. Thus a click of the type described has the highest potential for precision human echolocation.

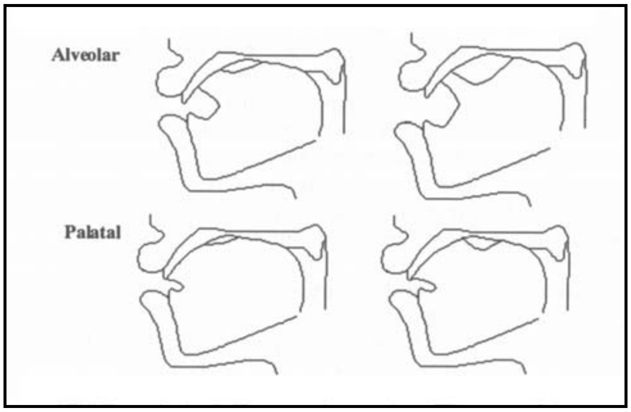

Figure R.9.1-2 The pre-release and post release position of the tongue during click formation. From Ladefoged & Maddieson, 1996.

---

[7] Ladefoged, P. & Maddieson, I. (1996) The Sounds of the World's Languages. Oxford, Blackwell

R.9.2 Examples of human clicks

Ladefoged & Maddieson have devoted their Chapter 8 to the characteristics of clicks. There are five places of articulation at which click consonants occur. In IPA, a click is symbolized by placing the assigned symbol for the place of click articulation to the left of a symbol for a non-click sound at the same place of articulation.

The easiest clicks for English speakers are the dental clicks written with a single pipe, |. They are all sharp (high-pitched) squeaky sounds made by sucking on the front teeth. A simple dental click is used in English to express pity or to shame someone, and sometimes to call an animal, and is written *tsk!*

Next most familiar to English speakers are the lateral clicks written with a double pipe, ||. They are also squeaky sounds, though less sharp than |, made by sucking on the molars on either side (or both sides) of the mouth. A simple lateral click is made in English to get a horse moving, and is conventionally written *tchick!*

Then there are the bilabial clicks, written with a bull's eye, ⊙. These are lip-smacking sounds, but without the pursing of the lips found in a kiss.

The above clicks sound like affricates, in that they involve a lot of friction. The other two families are more abrupt sounds that do not have this friction.

With the alveolar clicks, written with an exclamation mark, !, the tip of the tongue is pulled down abruptly and forcefully from the roof of the mouth, sometimes using a lot of jaw motion, and making a hollow *pop!* like a cork being pulled from an empty bottle. These sounds can be quite loud.

Finally, the palatal clicks, ?, are made with a flat tongue, and are softer popping sounds than the ! clicks.

The forward alveolar clicks are more useful in echolocation, particularly at close range, because they are propagated straight forward and contribute to a useful stereophonic perception in conjunction with the bilateral ears. While the lateral clicks may be louder, they tend to be projected away from the side of the face.

Figure R.9.2-1, also from Ladefoged & Maddieson, shows two clicks based on the above terminology. The palatal click shows a briefer duration and a less well defined sinusoidal structure. The alveolar click shows a nearly ideal gated sinusoidal form. Both of these clicks involve pressure changes during the click on the order of 10 cm of H_2O (about 1/100th of an atmosphere). The describe the rising characteristic of the click as the crescendo and the falling characteristic as the decrescendo. They note the alveolar click has a fundamental frequency near 1200 Hz while the palatal click has a fundamental frequency near 3000 Hz (reflecting the fact the air filled cavity is smaller in this case).

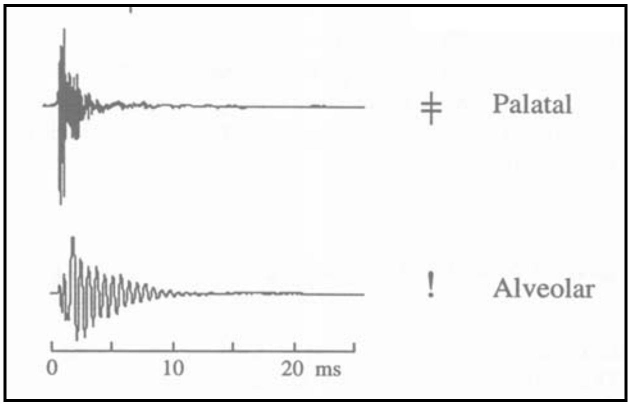

Figure R.9.2-1 A palatal and an alveolar click showing their individual forms and the standard symbol used to describe them. From Ladefoged & Maddieson, 1996.

What clicks sound like

For several sound samples see bilabial click, dental click, lateral click, palatal click, and alveolar click in Wikipedia.

The airstream

The essence of a click is an ingressive airstream mechanism. Although not often discussed, the forward alveolar click is made with the edge of the tongue contacting the palette all around its edge inside of the teeth and blocking off the windpipe. When the tip of the tongue is withdrawn quickly, a volume at low pressure is formed between the palette and the tongue. This volume of low pressure is projected outward as air from around this volume attempts to equalize the pressure. The precise frequency and duration of the projected sound depends on the precise shape of the lips and cheek tissue.

The linguists have defined the forward alveolar click much more precisely. The forward place of articulation is alveolar, which means it is articulated with the tip of the tongue against the alveolar ridge. They also speak of the front articulation as being coronal. The rear articulation has traditionally been thought to be velar or, again more rarely, uvular.

Since in at least some languages the closure of the rear of the mouth is not velar, some phoneticians have recently come to prefer the term *lingual* (made with the tongue) as being more accurate for this airstream mechanism than *velaric* (made with the velum).

For the clicks of interest here, there is only a single release burst, that of the forward click release, and the release of the rear articulation isn't audible.

Ladefoged has provided an X-ray sequence of the forward alveolar click in action as a QuickTime movie with sound.
http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/vowels/chapter13/movie.html

This page from the University of Stuttgart has a nice sound file of clicks:
http://www.ims.uni-stuttgart.de/phonetik/EGG/page5.htm

The click at 6.02 seconds consists of 4 cycles of a relatively pure fundamental at 666 Hz plus harmonic content. The leading edge of the pulse is negative going as recorded and lasts for about 300 microseconds. It is not clear whether the human ear can sense a leading edge extending for only 300 microseconds. Alternately, the Outer Hair Cells sensitive to 666 kHz may integrate the energy at that frequency over the complete interval of about six milliseconds. Such a received click could be perceived without interference from the outgoing click at ranges exceeding 100 cm. If a shorter series of sine waves could be generated, the minimum non-interference range would be less. These numbers assume compensation for any delays associated with the neural system, as commonly found in bats and probably found, though less clearly documented, in dolphins. This calculated minimum range appears compatible with the minimum range used by Ben Underwood in the video. **Figure R.9.2-2** shows this click using Raven Lite 1.0 software from Cornell University. The spectrographic presentation is quite dependent on the setting of the spectrogram "sharpness" control in this software (which apparently controls the temporal width of the Fast Fourier Window used to create the spectrogram.

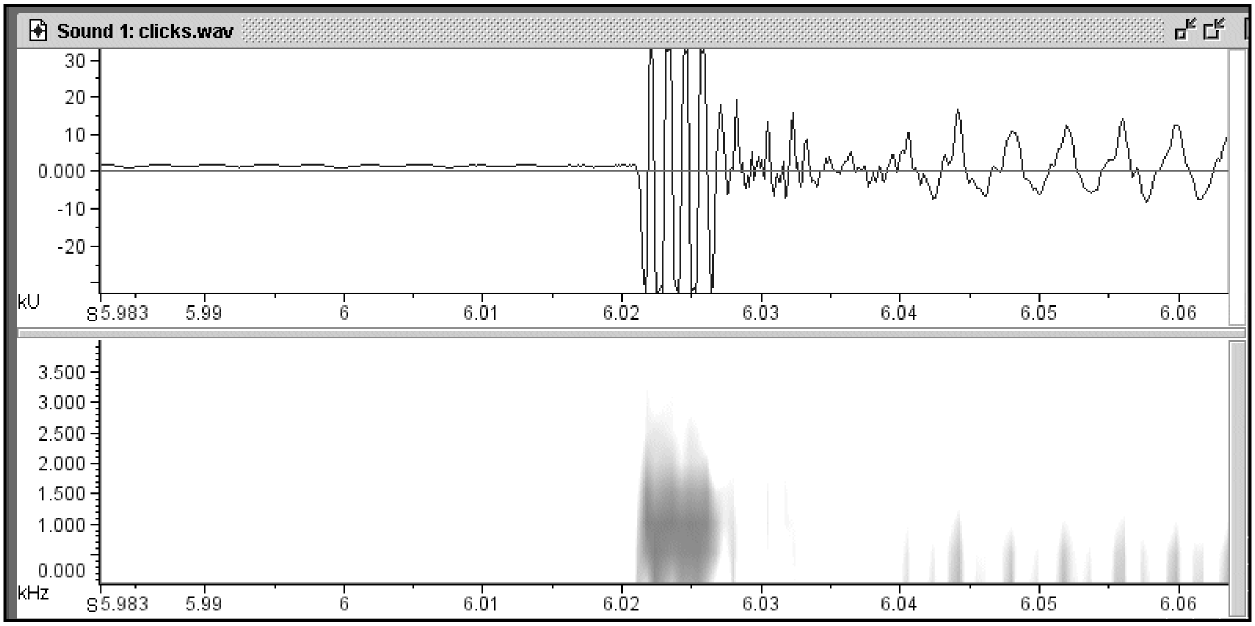

Figure R.9.2-2 A human click at the front of an African language word. The fundamental is at about 666 Hertz. From University of Stuttgart website.

The maximum acoustic range of human echolocation is dependent almost entirely on how intense a signal can be projected from the mouth. In optimum locations, echos can be perceived over distances of miles. However, this requires very loud emissions (not compatible with forward aveolar clicks) and very large reflecting surfaces under quiet conditions. Under more common conditions, it appears Ben Underwood can echo locate, at least in range, out to at least 10 feet. [xxx confirm this estimate]

A similar set of clicks from Ladefoged is available at http://hctv.humnet.ucla.edu/departments/linguistics/VowelsandConsonants/vowels/chapter13/nama.html Click number 15 on this page, the unvoiced click in the Nama word for belt, has a very pure but decaying 4 cycle tone at 666 Hz occurring 0.77 seconds into the recording. **Figure R.9.2-3** shows this waveform as displayed using Raven Lite 1.0 software from Cornell University..

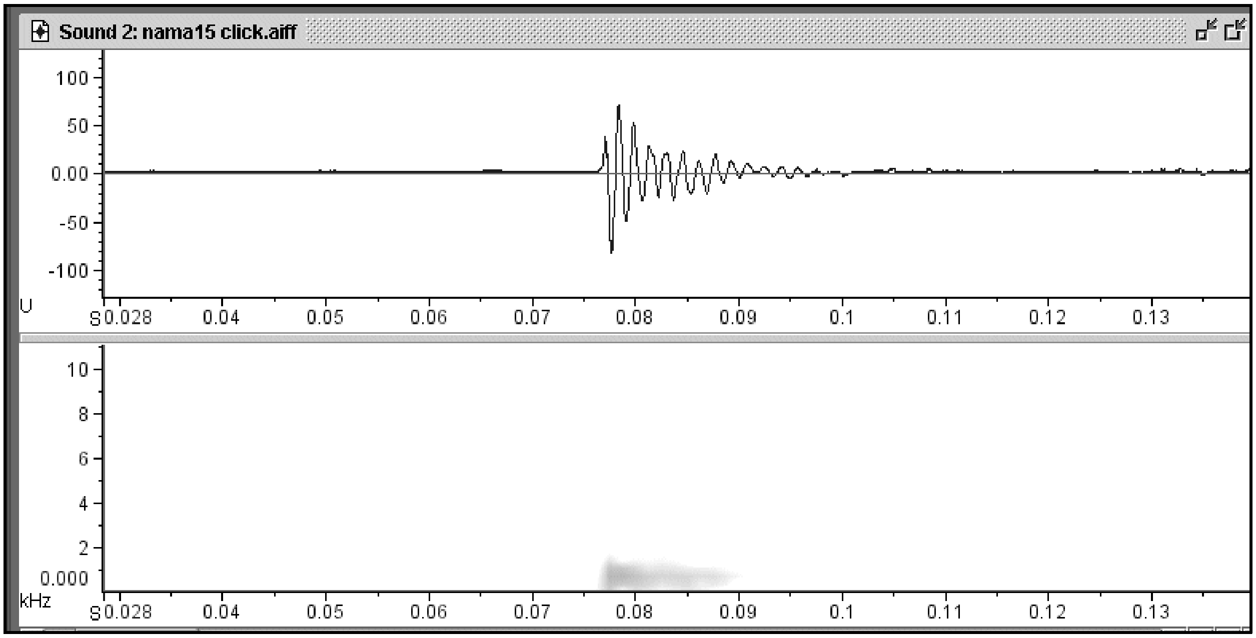

Figure R.9.2-3 The click as the first phoneme of the word “belt” in Nama. The tone is quite pure for a speech tone. It has a fundamental near 666 Hertz. From Ladefoged on UCLA website.

There are many other click waveforms available on the Internet but their provenances are poor to non-existent.

R.10 Support, Plasticity of the brain

“Even Old Brains Seem Flexible Enough To Enjoy a Workout” by Sharon Begley. Wall Street Journal. (Eastern edition). New York, N.Y.: Dec 2, 2005. pg. B.1

IF "TRUTH IN LABELING" laws applied to the three pounds of tofu-soft tissue inside your skull, the brain would be in big trouble.

Open any lavishly illustrated brain book to the diagram showing which region of the brain handles which tasks. There are strips that process touch, areas that handle sounds, even clusters of neurons that do math, get jokes and match verbs to nouns. But a growing chorus of researchers is saying, not so fast. These days, the brain's zoning map -- with different neighborhoods assigned different functions -- is looking as malleable as putty. Evidence of this "plasticity" has been piling up for more than a decade, but now neuroscientists are seeing that it is more radical than they thought, and that it lasts well into adulthood. Yes, you can teach an old brain new tricks. Take the visual cortex, which turns out to be quite a job hopper. In 1996, scientists using fMRI to peer inside the brains of blind people reading Braille found that the visual cortex processes the sense of touch.This big hunk of neural space (visual regions take up 35% of the brain, and 35% of a brain is a terrible thing to waste) noticed that no signals were arriving from the eyes, and looked around for other employment possibilities. With streams of input arriving from the fingertips, the opportunity was obvious.

PEOPLE WHO BECAME blind later in life didn't show this "cross modal" plasticity, suggesting that old brains can't change jobs. But many of those late-blind people lost their sight slowly, to diabetes, for instance. This may be too slow for the visual cortex to notice. When blindness comes on suddenly, the brain is remarkably nimble even in adulthood. A few years ago Alvaro Pascual-Leone of Harvard Medical School and colleagues blindfolded healthy, sighted adults for a week.

Every day, the recruits studied Braille. After mere days, their visual cortex was processing touch. This job switch happened too quickly to reflect new neuronal connections from, say, the fingers. Instead, those connections must have always been there, Dr. Pascual-Leone suspects, and become "unmasked" only when needed. That suggests that the visual cortex is misnamed. It doesn't see, necessarily, but makes spatial discriminations. "It's easier to do this with vision, but if there is no visual input it can rope in the next-best things, like feeling or hearing," he says. Indeed, in congenitally blind people the visual cortex also localizes sounds, figuring out where a noise came from. The visual cortex can also become a linguist. Harvard's Amir Amedi and colleagues recently found that people blind from birth seem to use their visual cortex to, of all things, generate verbs when an experimenter says a noun. "Apple" elicits "eat," and "piano" brings "play." But if researchers temporarily knock out the visual cortex with a magnetic pulse, the blind come up with semantic nonsense, such as "sit" for "apple."

The malleability of the brain well into adulthood can be a cause of both concern and optimism. The down side is that artificial vision, using tiny cameras to capture images and send them to the visual cortex, may be a pipe dream. Unless it's done soon after birth, which may not be practical, those images will be landing in a visual cortex that has moved on to other jobs, and the signals will not produce sight. THE UP SIDE is that old brains are continuing to learn. At last month's annual meeting of the Society for Neuroscience, researchers presented the results of a study in which elderly volunteers, 61 to 94 years old, underwent eight weeks of computer-based training to improve the brain's ability to discriminate the sounds of speech. "In the elderly, there is good evidence that the brain's representation of speech becomes noisier and degraded, which is why some have trouble understanding muffled speech or the speech of young kids," says Michael Merzenich, University of California, San Francisco. "If you have trouble processing fast phonemes, the information fed into memory is crummy." Many dyslexic children have the same speech-processing deficit. Prof. Merzenich and colleagues have shown that retraining the kids' auditory cortex through specially constructed language input improves their reading ability.

With similar retraining, the older brains, too, processed speech and remembered things better. "The majority improved 10 or more years in neurocognitive status, so 80-year-olds had the memories of 70-year olds," says Prof. Merzenich. "With more training, I expect we could get it to 25 years." He foresees a day when the discoveries of neuroplasticity will usher in "a new brain- fitness culture," reflecting "an understanding that you need to exercise your brain as you exercise your body." Crossword puzzles, bridge, reading and other activities aren't enough, though, especially if they've become routine. One of the most robust findings in neuroplasticity is that operating on autopilot doesn't help. Only mental tasks that you focus on intently can produce the physical changes that let old neurons learn new tricks.

Table of Contents

List of Figures