Journal of Communication Disorders, Deaf Studies & Hearing Aids

James, Commun Disord Deaf Stud Hearing Aids 2019, 6:1
DOI: 10.4172/2375-4427.1000188

Research Article

Open Access

Tactile Phonetics and Digital Impressions to Assist the Hearing and Visual Impaired

Samuel Martyn James*

Sharon Systems, Ashok Nagar, Chennai, Tamil Nadu, India

***Corresponding author:** James SM, Sharon Systems, Ashok Nagar, Chennai, Tamil Nadu-600083, India, Tel: +917305873254; E-mail: samuel@sharonsys.in

Received date: May 01, 2018; **Accepted date:** March 23, 2019; **Published date:** March 30, 2019

Abstract

"Touch to Feel" is a method consistent with natural behavior and responses. It is commonly prevalent in the living, human and animals alike. With a digital conversion of the signals, more accurate and compressed data can be transmitted to the recipient. A small test was conducted on children with hearing disability and without any training or understanding, in front of public; they were able to identify the digital conversion of phonetic signals presented in a tactile format and matched it easily with the corresponding sentences and words.

The purpose of this project was to find means to alleviate the struggle of the disabled to connect with the society with clear understanding and find a normal life in their current environment. Sign language does not give the hearing impaired to connect with all neither do the visually impaired are clearly able to perceive their surroundings.

A new method of impression for digital data would help both kinds of disability to efficiently address the issue and improve their perception and understanding of their environment, people and knowledge acquisition methods, so that they a fair opportunity to lead a normal life.

Keywords: Disability; Hearing impairment; Tactile; Wearable; Speech impression

Introduction

39 million fully blind and 360 million hearing impaired people are living alongside us according to WHO estimates [1]. Technology has taken great leaps in every major field and amazing things are being done through those inventions. Another marvelous thing that could alleviate the sufferings of these people could be achieved with some effort just using the available proven technologies. Helen Keller once stated that while blindness separates people from things, deafness separates people from people. What a sad and incontrovertible truth. A challenging thought that should propel every innovator to think and respond.

It may be mistakenly construed that acknowledging their rights and providing means to have sign language interpreters at news and conferences, and teaching schools is sufficient enough to make this go away. The crux of the problem is that these people are disabled and disconnected. They have no means to communicate with anyone who are not bilingual in the sign and spoken languages. They do spend all their time separated and alone, perhaps wondering what is happening in this beautiful environment around them.

Watching or working with them often makes one realize how frustrated they are. They expect everyone to understand their sign language; and if we don't it make them feel inadequate, irritated and even lost in a world full of wonderful people around them. Their sad plight is not limited to them alone. Considering their parents, siblings and friends, who struggle to get their thoughts into them and see a reciprocal gesture as they would normally expect from a child, brother or friend is also a sad situation. This problem has been in existence for a long, long time and did not emerge suddenly. There were many wonderful people who have addressed these issues and found some method to communicate with them. Sign language, Braille is notable accomplishments.

What we intend to achieve with this project would go beyond what is available in the current scenario. It is a tactile method that would communicate data in a way that is easily and readily comprehended by the user; use it to enhance their perception and respond correctly and adequately. By this the disability of the person would considerably be reduced and make it much easier for them to get closer to their environment and the people around them. As an example we have proved that it is possible for young children, even without any training or knowledge of the concept were able to identify tactile patterns correctly and immediately. Another interesting factor we noted was that not a single normal person was able recognizing the patterns correctly. This indeed goes to prove beyond doubt that the perceptual ability of the disabled is extensive and accurate. Using this as our base we have embarked on this project and a certain amount of confidence that it would make a lasting contribution to remedy the causes of the disability.

Rhema-tactile device for the hearing impaired

The most important principle factored into this project is cost, ease of use and readily available technology. Even if a single factor of the above had been difficult to achieve, the project would have made little sense to the needy and a dream for the future.

The three important steps to make this project work were identified as follows:

1. Audio processing, word recognition and linguistic models (corpus).

2. Pattern generation and matching words to patterns.

3. Converting patterns to digital data, transfer to serial and operate solenoids according to patterns.

There were other concepts such as audio filtering, voice recognition, direction etc. But these were left for future Research & Development, even though much of them are available as open source material.

For Audio Processing and word recognition and English Language (US) corpus we selected the open source software Sphinx provided by the Carnegie Mellon University & Wall Street Journal (CMU-WSJ). Thanks to their benevolence, using Sphinx, it was easily adapted to train voices in the Tamil and English spoken by the local dialect. A small Tamil language corpus was built and trained to identify the words and sentences spoken through a microphone into the system.

The first step taken was to see if the children would be interested in a new system other than the sign and lip reading methods. Software was designed to present the audio waveform of a spoken word, simple and in daily use. Since most of the speech impaired pupils had good vocal chords but could not use them, it was presumed that with a audio waveform, these children will be enthused to modulate their voice to the pattern on the screen.

The results were encouraging. In 30 minutes a small boy was able to modulate and train his vocal chords to pronounce 3 words, "amma, idly, dosai" with a match result of >80%. The system was designed to portray a picture and text of the word pronounced correctly. This encouraged the other pupils and because we didn't have anything else to offer, withdrew to continue in the project further.

The next step was to find easily recognizable patterns for the words that are spoken in any language. The system had to be language independent as the motive was to generate a system that could be used by all in any language or dialect setup. The age old versatile system that had served thousands if not millions in the pre audio recorder era, the Pitman's short hand system seemed to be the logical choice. It was phonetic based, with limited symbols and easy to adapt, it was a preferred choice as a first module. Later it was found that some of the symbols and grammatical rules were quite complicated; and in those instances the Hebrew Language rules and consonant patterns were adapted into the system.

The next task was to assign digital patterns to the phonetic and consonant symbols that were adapted into the system. A 4 x 4 array for the hardware module was designed and the serial inputs given to the controller activated the LEDs accordingly. Having attained the goal, it was then decided to follow through with solenoids instead of the LEDs, which would give a tactile impression that could be felt by the user without having to see it.

The device seemed large and heavy. There was no way anyone could feel comfortable wearing it on their waist and the power consumption would have been prohibitive. It was then decided to reduce the size of the array to 3 x 4 and two types of solenoids, one a mini version weighing 10 g for light duty and the other weighing about 30 g for heavy duty usage was brought in. The heavy duty one was embedded into a table-top suitable for use in a class room.

It was wired to a system through hardware especially designed for the purpose. The system operated by a voice input into the microphone attached to the system. The audio was recognized and converted into a pattern. The solenoids were activated according to the assigned values, and the pupil was able to pick the pattern by identifying the related sentence or work displayed on a chart.

The one factor that impressed all who were present was that the children did not require any training or instruction before the demo and they continued with keen interest and excitement. The intent to produce a cost effective, easy to use and linguistic inter-dependent had been made possible by this project.

Method

Sphinx audio processing

The first major task, to convert audio into recognizable phonetic syllables was made easy by the CMU-WSJ benevolence. It would have been a daunting task for anyone, especially from the third world, without the resources and funds available in the US. Thanks to the provision of the Sphinx codes, the task of separating words from a sentence by dissecting the audio waveform from a recorded speech in Tamil was achieved. The process was not as easy as it seems, and would have been impossible for me without the assistance of Clarence Fernando, my colleague who assisted me in the technical implementation side of the project.

Once the words were dissected into phonemes through the Sphinx software it was then stored in a database along with the audio files indexed to the phonetic words.

Tactile symbols generation

The symbols were designed for the phonetic words and then were stored in the system. An elaborate study was made based on the phonetic structure of the words, Pitman's short hand and the Hebrew Language.

Phonemes are classified as 44 in number with 18 consonants. The English language alphabets were 26 in number. Whereas Pitmans had 27 and the Hebrew language had 32 consonant symbols. Combining the three systems, we stucked with the phoneme structure of 18 consonants and added two more symbols for vowels, one for the preceding vowel before a word and one for the following vowel after a word.

The Hebrew language had with stood time and usage without any vowels and the Bible is the most accurate original version that has remained the same over 6 millennium. With conscious efforts vowels can be built into a word and understood as the speech goes by. It can save a lot of complications and confusions, while trying to tactile every utterance. It also helped in keeping the tactile impressions limited to 20 symbols, which is a very small number for any child or adult to learn and master within weeks.

The 2000 words published in the General Service List contains a phoneme pattern that matches other forms of materials collected for analyzing the common letters that appeared in each of them (Figure 1) [2].

It will be seen that the alphabet pattern remains the same almost. Using this as a criterion it was decided to represent the bottom 8 alphabets in a different manner, thus keeping only 18 patterns based on the phoneme consonants [3]. Adding 2 for the preceding and following vowel the goal was achieved to limit tactile impressions to 20 [4].

Citation: James SM (2019) Tactile Phonetics and Digital Impressions to Assist the Hearing and Visual Impaired. Commun Disord Deaf Stud Hearing Aids 6: 188. doi:10.4172/2375-4427.1000188

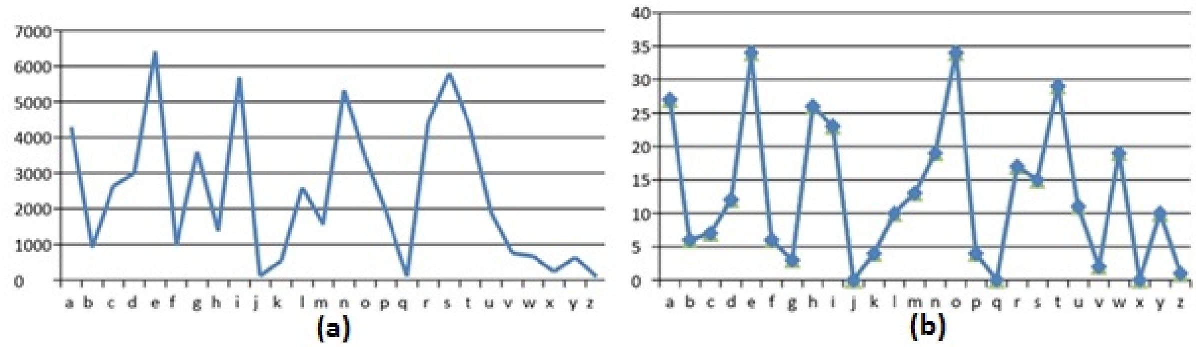

Figure 1: (a) No. of times each alphabet is used in the top 2000 general service list. **(b)** Compare this with the top 100 words used in the English Language.

Audio to word recognition

The next step was to implement a system where the pupil will learn to recognize the audio patterns from a wave form displayed on the screen. Those with good vocal chords and unable to use them because of the hearing disability were given an opportunity to see and practice their voice. The recorded words stored in the system were used for the application. The data base had a picture and a text file indexed to the word.

When a word from the data base was pulled by a mouse click, the wave pattern for the audio and a text was made visible on the screen. A computer with audio recording software was integrated with a modified version of audacity to display a wave form pattern for the recorded voice. It was also programmed to record voice and display the test audio live on screen. Tools were given to move and expand the waveform to match with the original and analyze the difference.

When the voice pattern matched above 80% of the original recording, a picture and text were displayed alongside to indicate there was a recognizable match. This would give an impetus to the children to try more words and understand the logic behind the program.

Word to sentence formation

Learning a language in the tactile format is not an easy task. A language consists of grammatical rules, localization and dialects. Expounding all these in a very clear and crisp output is certainly bound to give errors and inaccuracies. Here it seemed logical to use the Pareto 80/20 law for the purpose of forming sentences. 20 percent of the commonly used words cover 80 percent of the sentences used in day to day affairs. Specialized conversations such as technological, spiritual, political, economical etc. subjects have words that not commonly used. These are however words that are learnt by association. One makes a conscious effort to understand these words, idioms and phrases. They can however be interpreted as a tactile impression and with practice such words will become common to the user who will be trained in a particular vocation.

During the training period the trainees will be taught to recognize a word pattern based on a single stroke/pattern.

Example:

ch+r = chair

g+ch+r=get me a chair (to be polite they will use 'please' as prefix)

ch+l=chill

ch+l+d=child/chilled

ch+l+r+n = children

The advance version with lesser symbols would evolve itself, when the children learn to interpret symbols and associate it with sentences they already know. It is the same manner in which our brain works with our eyes, while reading.

Adaptive dictionary-based compression

The visually impaired already have a form of tactile experience with Braille. The Rhema system can be easily adapted by them, as it contains only 20 phonetic characters, combining numerals and words. Whereas the Braille system has 64 signs created from 6 dots.

The inclusion of the adaptive dictionary-based compression algorithm would richly enhance the competence of the disabled person. By translating data from a video or audio input into recognizable form and storing them into a dictionary as the user gets to use the system, recognizable impressions will be conveyed to the user quickly and efficiently [5]. The size of the data being communicated locally will require much less bandwidth and can be played back again and again.

By using this method a visually impaired person can recognize the environment, people from known and unknown, moving and static objects etc.

When a known person is in the visual range of the video equipment, the name of the person will be converted to a tactile format, which could be felt by a visually impaired person. The familiar environment will indicate a change in object position or missing and replaced. Likewise an image will be analyzed and the pixel data in the compressed format will be relayed as a tactile impression. With a good and powerful system, it is not impossible to get them to paint an image by reading the pixels data of the image and reproducing it on a canvass.

A hearing impaired person can be given a clear audio filtered according to his needs such as near and far voices, known and accustomed from the strange noises in the surrounding area; know the

Citation: James SM (2019) Tactile Phonetics and Digital Impressions to Assist the Hearing and Visual Impaired. Commun Disord Deaf Stud Hearing Aids 6: 188. doi:10.4172/2375-4427.1000188

direction from where being addressed or even tune to a speaker on a stage with little noise distortion [6].

Recording frequently spoken words in the dictionary, according to a profession, branch of study, interest or hobby and above all known and close relatives is completely possible and can be converted into a tactile impression with ease.

Dictionary-based compression

Example: Consider the random house dictionary of the English Language, Second edition, Unabridged. Using this dictionary, the string: A good example of how dictionary based compression works can be coded as: 1/1 822/3 674/4 1343/60 928/75 550/32 173/46 421/2 Coding: Uses the dictionary as a simple lookup table each word is coded as x/y, where, x gives the page in the dictionary and y gives the number of the word on that page. The dictionary has 2,200 pages with less than 256 entries per page: Therefore x requires 12 bits and y requires 8 bits, i.e., 20 bits per word (2.5 bytes per word). Using ASCII (American Standard Code for Information Interchange) coding the above string requires 48 bytes, whereas our encoding requires only 20 (<-2.5*8) bytes: 50% compression [8].

By storing data in this format in the Arduino board and PC or tab, the recognition speed and TX speed will be increased considerably, giving a more accurate translation of the speech data.

The possibilities are extensive; they are all proven technologies, economical to develop and easy to spread with an open source administration. The benefits of the system can be brought to the needy all over the world in a short span of time. There is no possibility for a physical hurt from electrical or other kinds of danger. The gains are far outreaching for them.

Hardware

The sentences formed by audio or input texts are then being converted into serial data in the normal method using an Arduino board linked to a PC. The Arduino board was picked because it has the capability to work independently and with the required 12 ports GPIO (General Purpose Input Output) for serial data TX. It could be hardwired to a PC or connected to an Android device through Bluetooth. It is also low priced and easily available in the market.

The symbols matching the word or sentence are transferred to the Arduino board and in turn the serial data were transferred to the solenoid driver PCB. The tactile device is controlled by the driver PCB and is activated according to the data received by the controller from the Arduino board (Figure 2). The whole process is completed at the rate of <30 ms per bisyllable word. This includes the time gap between syllables. We have provided a short time gap of 2 ms between syllables and a 5 ms gap between words. This enables the user to read and understand the impression clearly. More than 100 wpm capability has been achieved in this process [7].

In a basic (beginner) version the Rhema system can achieve a rate of 100 to 120 words per minute (wpm). In the advance method where an adapative dictionary-based compression method is used to form sentences and special words, a speed of above 150 to 180 wpm is expected to be achieved.

Three types of solenoids were used with varying sizes. One large and one smaller heavy duty solenoids with working on 1 to 2 W, 3.5–5 VDC (Volts Direct Current) ranging from sizes 10 mm to 20 mm square on the top and base.

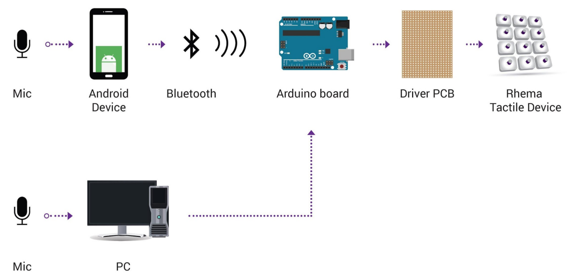

Figure 2: Assembly of the device.

Both the heavy duty ones were for the table top version, where the use were for various pupils to be trained on the system. It was wired to a PC based system and networked with microphones and speakers, so that each individual received the same power and signal.

The smaller solenoid version was intended for the body worn variation of the device, was wired to the Arduino board and controller driver with a power supply battery unit. It weighed about 500 g. The communication with the PC or the smart phone was through blue tooth.

Responding Methods

The above system enables the disabled to understand the communications addressed to them. But it is only half duplex; that is one way. Only when the user gets to respond adequately to the communication will the process of a 2 x half duplex (RX/TX) communication be completed. Keeping this in mind a system was developed to allow the user to quickly respond using an android tab or smart phone.

The android application contains a method to record and store audio or text messages that are standard or able to retrieve from the Adaptive Dictionary-based Compression files by pressing a button on the android device screen. These data are indexed in an alphabetical order and the user can access the stored data by pressing the button and choosing the relevant response. The chosen sentence will then be played back on the speaker, or even transmitted over the phone to a remote listener.

Now the user can feel and understand what is being communicated to them, and also respond to a conversation, discussion or lecture properly, in real time and also know what they have just said.

Conclusion

Rhema-Tactile device for the hearing impaired is a novel method aimed to relieve some of the difficulties faced by the hearing and the visually impaired. It addresses the concerns of the hearing impaired as a means of communication: in a half-duplex RX/TX mode.

To receive communication a conversion to tactile process is made using simple and readily available technology and the proven Pitman's and Hebrew system and rules.

To transmit a response an easy to follow method was incorporated using available technology to record, store and retrieve words, sentences and phrases as are common to the user. The response could be dynamically or selectively chosen and sent to an audio player or phone outgoing channel.

The visually impaired would now be able to benefit with shorter impressions and faster reading. They can also benefit from video capture and analysis method, by being able to recognize environment, people, images etc. converted into readable tactile impressions.

By using available and proven technology, a cost effective system that comprises of software, a tab and a tactile device would not exceed $ 200 at the production level.

It has the advantage to be able to adapt to any language with the relevant language corpus.

Anyone with little experience can modify the system to suit their requirements and it can be distributed globally in a localized format easily and fast. The benefits of the system could be made available to all sooner than later.

References

1. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment
2. https://en.wikipedia.org/wiki/General_Service_List
3. http://www.duboislc.org/ED-Watch/Words/1-100.html
4. http://englishharmony.com/
5. Nelson M, Gailly JL (1995) The Data Compression Book. (2nd Edtn), MIS:Press New York, USA.
6. http://www.drdobbs.com/architecture-and-design/lzw-data-compression/184408217
7. http://collaboration.cmc.ec.gc.ca/science/rpn/biblio/ddj/Website/articles/DDJ/1997/9718/9718d/9718d.htm
8. https://users.ece.utexas.edu/~ryerraballi/MSB/pdfs/M2L3.pdf