



## **AUTOMATIC SPEECH RECOGNITION FOR THE HEARING IMPAIRED USING AUTOMATIC ROBOT PROCESSING**

**D Rajkumar, Research Scholar, Dept. of Computer Science, Himalayan University**

**Dr. Manish Saxena, Professor, Dept. of Computer Science, Himalayan University**

### **ABSTRACT**

Speech is a very important tool for human interaction. ASR refers to the process of giving computer system intelligence through the use of computational algorithms that can identify spoken language and translate it into text or produce a control action. ASR is a method through which a machine may recognize the human speech used as input and carry out the specified task. The main responsibility of such a system is to create a speech recognition model. However, unlike people, computers do not possess the intelligence to understand speech. Humans are capable of distinguishing the sound of interest among a variety of concurrently audible sounds. A computer-based approach, on the other hand, will classify other noises that are not of interest as noise. Therefore, advancements are crucial if one is to create a reliable computer system. Speech encoding, speech segmentation, and noise removal in speech make up speech pre-processing. In the feature extraction stage, a computational technique is used to identify the distinguishing characteristics of the voice sample. Different supervised or unsupervised speech classification algorithms are used in the classification stage. ASR systems are divided into isolated word, connected word, continuous speech, and spontaneous SR systems depending on their kind. According to speaker mode there are three types of ASR systems: speaker dependent, speaker independent, and speaker adaptive. Features in speech technology are measurable qualities that are present in each sample.

***KEYWORDS: Speech, Reliable, Recognition, Audible, Measurable***



## 1. INTRODUCTION

The process of auditory perception of sound is referred to as hearing. Normal hearing is dependent on the structural and functional integrity of the auditory system. Hearing loss is brought on by problems with the auditory system. Any level of impairment in the capacity to perceive sound is referred to as hearing loss. The most prevalent sensory deficit in the human population is hearing loss, which affects more than 360 million individuals worldwide. 32 million (9%), of these are children, while 328 million (91%) are adults (183 million men, 145 million women). According to the World Health Organization (2017), Sub-Saharan Africa, Asia Pacific, and South Asia have relatively high rates of hearing loss. According to Ciorba, hearing loss can also significantly damage a person's functional, social, economic, and emotional status. This can have an impact on the person's overall quality of life. By providing an adequate auditory rehabilitation program for those who have hearing impairment, the effects of hearing loss can be lessened.

The term "device, procedure, information, interaction, or therapy" used to describe aural rehabilitation is "anything that lessens the communicative and psychosocial consequences of a hearing impairment". Aural rehabilitation aims to improve or restore a person's limited involvement in activities due to a hearing impairment. It may also be meant to encourage communication partners who participate in activities with people who have hearing loss. According to Boothroyd, aural rehabilitation consists of four main components that can be categorized as follows: perceptual training (to target activity, by enhancing the learning opportunities provided by everyday communication), instruction (to increase the probability of a positive outcome from perceptual training), and counseling (to target issues of participation and quality of life that result from residual hearing loss). For successful results in auditory rehabilitation programs for people with hearing loss, all of these steps should be followed.

Only the sensory management, which typically involves giving hearing aids or cochlear implants, has grown in significance in the practice of aural rehabilitation for older people with hearing loss. Audiologists frequently overlook other crucial procedures including instruction, training, and counseling. The fact that the amplification device does not completely restore

hearing is important to note. The use of an amplifier merely enhances audibility and may not correct issues with frequency resolution or lost temporal cues, which results in a poor representation of all auditory cues. As a result, compared to people with normal hearing, the majority of hearing aid users experience auditory signals that are distinct and seem inferior.

### **1.1 NOISE WITH SPEECH**

Everyday communication takes place in situations where there are competing signals or background noise, not in an ideal acoustic environment. Even individuals with adequate hearing and cognitive abilities find it difficult to understand speech in noisy environments. In addition to this challenge, it becomes more difficult and demanding to separate complex signals like speech from the background noise if a person has hearing loss. The most frequently mentioned issue in people with sensory neural hearing loss is interpreting speech in the presence of noise, according to clinical practice and literature analysis.

According to reports, people who have hearing loss typically experience attenuation factor (lower audibility) and distortion factor (reduced Signal to Noise Ratio), both of which have a detrimental impact on how well speech is perceived. Simulating hearing loss in those with normal hearing can help to understand this. Individuals with sensory neural hearing loss perform poorly in the presence of background noise, despite the fact that the degree of hearing loss they experience is the same. As a result, those with hearing loss require better signal to noise ratios up to +8 dB in comparison to those with normal hearing. Due to impaired central or cognitive capacities brought on by age, or a combination of both challenges, difficulty understanding speech in the presence of noise is more noticeable in older individuals with hearing loss. The Hearing in Noise Test and the Quick Speech in Noise Test (Quick SIN) are two common tests used to measure speech comprehension in the presence of noise. Recent research typically uses the Quick SIN test among these since it is easier to use. Along with the results for voice detection, it also provides Signal to Noise Ratio Loss.

### **1.2 SCOPE OF THE RESEARCH**



Everything becomes digital in today's environment. Software and computers are ubiquitous in daily life. Computers can handle the majority of our tasks. Booking tickets, paying bills, booking hotels, shopping, and other tasks are all done using computers. A software engineer creates software that comprises of a collection of programs to carry out this operation. The ultimate implementation of this program can be tested at various phases of development.

The goal of the work is to design a method for reusing software parts to create improved software with interdependency characteristics of crucial elements. Software assets and components that may be reused have several benefits, including fewer code, faster development times, smaller file sizes, and other benefits. To achieve our goal, we apply a strategy in this work where we reuse software components from one piece of software in another.

The work's focus is on the software sector, where this reusability technique has a lot of potential applications. This can be used by enhancing the software's tracking system and then including the reusability strategy. As a result, we can create an improved project that is simple to manage and maintain.

## **2. REVIEW OF LITERATURE**

Karpagavalli et al. (2016) established a structure to aid in the correlation between speech in order to focus on personalized discourse affirmation. This research also investigates whether using a cross-breed strategy is more intelligent when planning components because it helps with feature extraction and vector quantization. Differentiating with various methods will be better and will demonstrate a range of preferences.

Combining all element extraction techniques yields the greatest results for a noisy environment, and this methodology's use of a variety of channels results in high precision. This methodology can be applied in a variety of flexible situations.

Preeti et al. (2013) examined fundamental methods for controlling discourse affirmation, and as a result, the outcomes showed higher exactness. Discourse outline, different types of discourse groups, feature extraction techniques, data and execution assessment are just a few of the

problems or issues that the arrangement of discourse affirmation structure would really like cautious considerations to the partner ASR system's acknowledgment performance evaluation should be based on a set of data that is wholly unrelated to the teaching corpus.

Santosh et al. (2010) provided a breakdown of a graph of the primary innovative choices made on the Feature course for creating systems in Marathi language-based human computer interfaces. The speech feature extraction has a severe classification disadvantage due to a reduction in the spatiality of the information vector while maintaining the sign's separation strength. We need to include discourse signal extraction because, as we collectively know from the fundamental arrangement of talker recognizable proof and confirmation framework, the amount of instruction and check vector required for the grouping disadvantage develops with the component of the given information.

### **3. RESEARCH METHODOLOGY**

The field of speech recognition has long been recognized as crucial to human computer interaction. A voice recognition system uses methods from a wide range of fields, including statistical sequence identification, communication theory, signal perception, and linguistics. The fundamental or main objective of research in voice processing is the human-machine interface. Inter-disciplinarity and the ability to address scientific problems are two of the most important goals of research into speech signal detection using machines. Hindi is our mother tongue, thus its local application is a significant and motivating factor in the choice of language for our appreciation framework, for instance. Real-time communication recognizers can be used extensively in a variety of settings, including simpler man-appliance signal, assistance for physically disabled and hearing-impaired individuals, telephone additional worker, and other man-appliance edge tasks. Due to its relatively straightforward methodology, the insulated word identification systems were among the earliest speech recognition systems produced. In this analysis, the idea and practice of Automatic Speech Recognition as a framework for Isolated Digit Recognition are discussed. The goal of automatic communication identification is to develop methods and systems for early computer systems that receive input from communication signals. Speech signal recognition evolves, such as the speaker recognition, isolated word

recognition, and speech signal to text conversion challenge. Users wanted the device to recognize their speech signals. By utilizing some basic elements, such as Wavelets, we have achieved significant progress in (ASR) designed for effectively-well defined presentations such as notation and average terminology operation doing duties in both contexts. In extremely loud environments, fiber optic communication is a cradle for big development. Our objective in this research is to acquire advanced techniques.

#### **4. RESULTS AND ANALYSIS**

Speech recognition is becoming a more useful invention in PC applications these days. There are many uses for conscious speech that are intuitive. Speech recognition technology must therefore advance in order to use this kind of simple communication strategy in the computer area. The computer must be programmed to receive voice input and then process it using different speech recognition programs to produce the desired output. Speech empowerment is the process of applying fitting analysis to recognize the speech sign in a group of words. This gives those who are less educated or have less computer experience a different and more effective approach to access systems when typing is challenging.

Since speech recognition doesn't require any grammar, process, or coding to access the commands, it is employed in numerous applications. It just takes spoken words as input and converts them into the proper text to complete the operation. With an interesting range of tasks, speech recognition technology offers an endless user base. The main goal of speech-enabling technology is to "listen", "identify", "understand", and "respond" to spoken information. It might function as a conduit between people and computers. Human-computer interface, etc.

In order to provide better execution in a short amount of time, we choose this Speech Recognition technique to be used in a Customer Care Organization throughout our research. They are responsible for responding to the inquiries that clients ask in the customer care organization. There are many other queries that could be asked, including ones about account details, server issues, server restarts, password resets, account locks, account balances, account transactions, last transactions, etc.

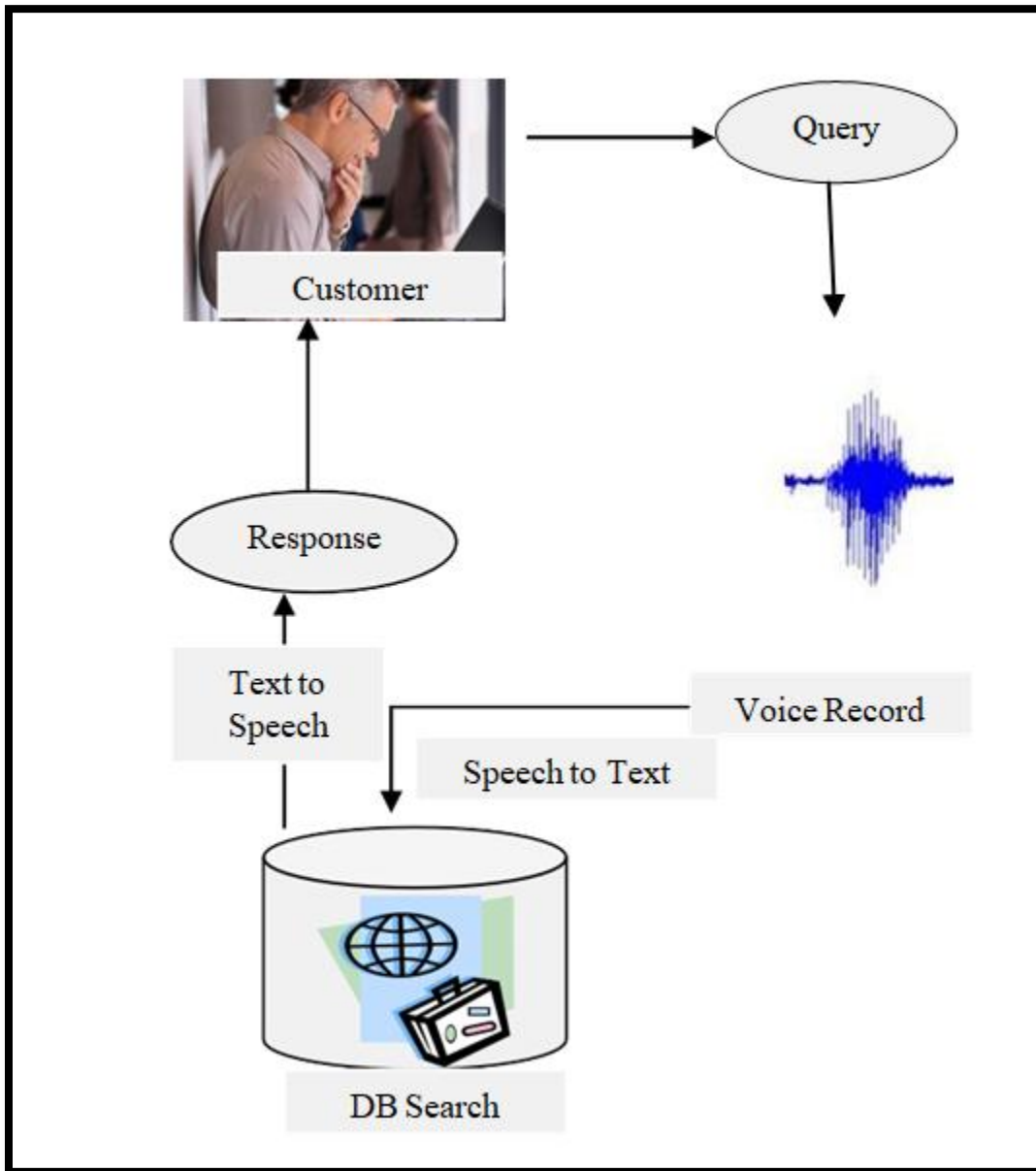


To increase customer loyalty and experience some pre-defined duties with reduced manual proficiency, we need to use the Speech Recognition innovation in the Customer Care Service area. Improved client relationships and better client administration are the main drivers. The act of providing outstanding administration in a timely manner while attending to client needs is known as client care service. Such a reaction requires the activation of an increasing amount of manual power. We must use a suitable strategy to achieve the goal of naturally responding to client needs while using speech recognition technology in order to avoid this problem. The development of a simple voice recognition system for customer care services using several technologies is depicted in this study.

In certain circumstances, the client may repeatedly ask questions, to which the official responds. Some of the questions might be asked again, and the official will inevitably have to respond to them. For this circumstance, it takes a remarkable amount of manual labor, 24 hours a day, seven days a week.

To overcome these circumstances, we must suggest a methodology that uses some pre-defined tasks to automate the process of supplying answers to inquiries. The Customer Care Server must be programmed with a set of instructions using the proper database keywords in order to accomplish this. The process can be carried out as follows:



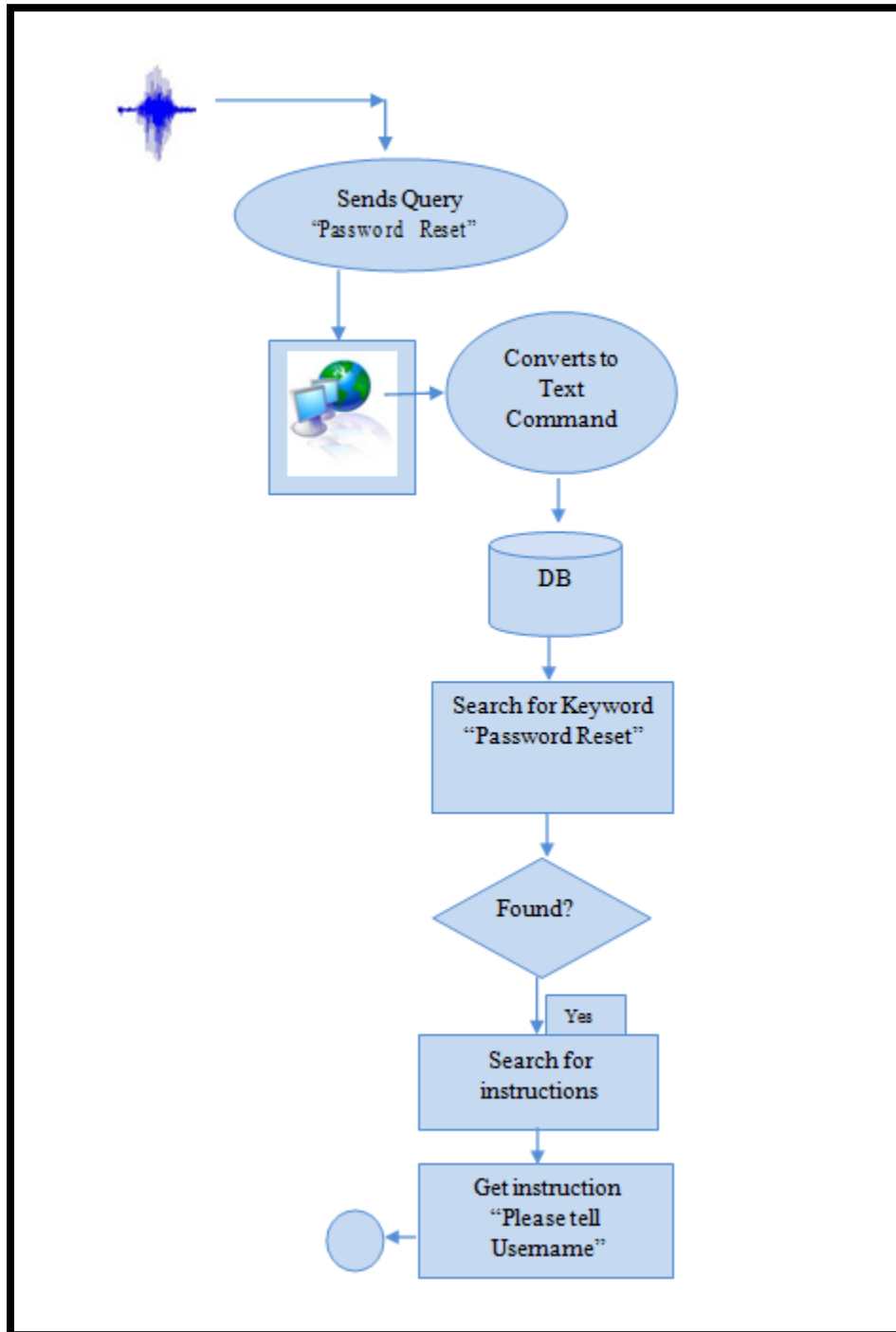


**FIGURE- 1PROCESSING THE CUSTOMER QUERY**

The server accepts the customer's voice request when they dial a number to make an inquiry, and records it. After receiving the input, the server transforms the voice into text and begins looking up the keyword in the text in the database. This database search yielded the proper query instructions from the database, and the response was sent to the customer. Thus, the speech



recognition technology, such as ROBOT Processing, has automatically processed the query. As a result, another name for our approach is Automatic ROBOT Processing (ARP).





If a consumer needs to reset their password, that falls under the "Password Reset" category. The customer initially phones the customer service line and uses the voice command "I Need Password Reset" to request a password reset. The server has noted down this command, which it then transliterates into text. The server then starts looking up the term "Password Reset" in the database. There are numerous instructions for the password reset category in the database that need to be programmed by the programmer to handle the request correctly.



## **FIGURE- 2 PROCESSING QUERY FOR PASSWORD RESET**

If the database has fields for "Please tell your Username" and "Please tell your Date of Birth" under the "password reset" category, the server will retrieve those instructions and convert them

to speech before transferring them to the client so they can speak their username and date of birth. As soon as the server receives the customer's details, it begins to transform them into text from speech and then verifies them. The server only permits a customer to reset their password if the data they supply matches that in the database. Otherwise, he or she must be turned down.

To make the process simpler and use less manual labor, the tiny predetermined tasks in customer care have been automated using speech recognition technology. Additionally, it lowers the expense of processing repetitive queries that have been predetermined. We named our methodology Automatic Robot Processing since the way we automatically process customer enquiries is similar to ROBOT type processing. The system identifies the voice data and employs appropriate methods to handle the consumer request. The system uses a count to show how the process is moving.

### **Procedure:**

#### **Start**

- Program Server with Voice Recognition Software
- Maintain Database with set of instructions for Customer Query Processing Receive the Customer Connection
- Get the Query as Voice Command
- Translate into Text Command
- Search in Database for related keywords
- Retrieve the appropriate instructions from the database
- Convert the instructions from text to speech commands
- Transfer the command to the Customer
- Retrieve the reply for the instruction from the customer as voice command
- Translate into Text Command
- Validate the details in the Database
- Send Response for the query if applicable

After doing extensive research, it was determined that Speech Recognition has carved out a place for itself in many spheres of life for addressing speech-oriented tasks. In the area of computers, many speech recognition framework methodologies are being looked at. The most manageable issue to move is this Speech Recognition. In this essay, we have attempted to provide a solution for the Customer Care Organization to automate a few pre-characterized tasks. We also offer an appropriate algorithm to handle this robot processing. In this instance, it appears simple to access the customer service center's inquiries without any manual resources.

## **5. CONCLUSION**

In contrast to other popular methodologies, this research's speech recognition additionally uses Mel Frequency Cepstrum Coefficients, which results in great accuracy and efficiency. Without significant intelligence and training, voice recognition by robots is difficult. The communication between humans and machines can be made simpler than it was in the past by creating such a controllable intelligence system. Speech recognition is a very difficult and complex field, and the more research it receives, the more success and development it will provide. Numerous advantages will result from new developments and discoveries in this field, improving people's quality of life generally and society as a whole.

## **REFERENCES**

1. Assmann, P., & Summerfield, Q. (2004). The perception of speech under adverse conditions Speech processing in the auditory system (pp. 231-308): Springer.
2. Avinash, M., Meti, R., & Kumar, U. (2010). Development of sentences for Quick Speech-in-Noise (Quick SIN) test in Kannada. Journal of Indian Speech and Hearing Association, 24(1), 59-65.
3. Baddeley, A. (1992). Working memory and conscious awareness Theories of memory (pp. 11-20): Lawrence Erlbaum Associates.
4. Bakin, J. S., & Weinberger, N. M. (1996). Induction of a physiological memory in the cerebral cortex by stimulation of the nucleus basalis. Proceedings of the National Academy of Sciences, 93(20), 11219-11224.

5. Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychology and aging*, 12(1), 12.
6. Barcroft, J., Sommers, M. S., Tye-Murray, N., Mauze, E., Schroy, C., & Spehar, B. (2011). Tailoring auditory training to patient needs with single and multiple talkers: transfer-appropriate gains on a four-choice discrimination test. *International Journal of Audiology*, 50(11), 802-808.
7. Basu, S., Prakash, H & Rajashekhar. B. (2008). Time compressed speech test in Kannada (Unpublished dissertation), University of Manipal, India.
8. Beasley, D. S., Forman, B. S., & Rintelmann, W. F. (1972). Perception of time-compressed CNC monosyllables by normal listeners. *Journal of Speech, Language, and Hearing Research*, 15, 340-350.
9. Beynon, G. J., Thornton, F. L., & Poole, C. (1997). A randomized, controlled trial of the efficacy of a communication course for first time hearing aid users. *British Journal of Audiology*, 31(5), 345-351.
10. Bode, D. L., & Oyer, H. J. (1970). Auditory training and speech discrimination. *Journal of Speech, Language, and Hearing Research*, 13(4), 839-855.
11. Boothroyd, A. (2007). Adult aural rehabilitation: What is it and does it work? *Trends in Amplification*, 11(2), 63-71.
12. Burk, M. H., & Humes, L. E. (2008). Effects of long-term training on aided speech-recognition performance in noise in older adults. *Journal of Speech, Language, and Hearing Research*, 51(3), 759-771.
13. Burk, M. H., Humes, L. E., Amos, N. E., & Strauser, L. E. (2006). Effect of training on word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners. *Ear Hearing*, 27(3), 263-278.
14. Burns, A., Brayne, C., & Folstein, M. (1998). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. M. Folstein, S. Folstein and P. McHugh, *Journal of Psychiatric Research* (1975) 12, 189-198. *International Journal of Geriatric Psychiatry*, 13(5), 285-294.



15. Census of India (2001). [www.censusindia.gov.in/2011-common/census\\_data\\_2001.html](http://www.censusindia.gov.in/2011-common/census_data_2001.html).
16. Chisolm, T. H., Abrams, H. B., & McArdle, R. (2004). Short- and Long-Term Outcomes of Adult Audiological Rehabilitation. *Ear Hearing*, 25(5), 464-477. doi: 10.1097/01.aud.0000145114.24651.4e
17. Chisolm, T. H., Abrams, H. B., McArdle, R., Wilson, R. H., & Doyle, P. J. (2005). The WHO-DAS II: psychometric properties in the measurement of functional health status in adults with acquired hearing loss. *Trends in Amplification*, 9(3), 111-126.
18. Ciorba, A., Bianchini, C., Pelucchi, S., & Pastore, A. (2012). The impact of hearing loss on the quality of life of elderly adults. *Clinical interventions in aging*, 7, 159-163.