A REVIEW ON: CLASSIFICATION OF SOUNDS BASED ON PHONETIC FEATURES

[1] Neha R Shetty, [2] Neha R Barki,[3]Nitish B V,[4]Mr.Mahesh kini M

[1] neharshetty7503@gmail.com, [2]nitishbv6@gmail.com, [3]nehabarki4@gmail.com

[4] maheshkini@aiet.org.in

ABSTRACT- **In linguistics, the categorization of sounds according to phonetic characteristics is a crucial field of research that sheds light on the processes behind human speech perception and production. With an emphasis on basic characteristics including place of articulation, mode of articulation, and voicing, this study aims to systematically classify speech sounds based on their phonetic characteristics. We seek to clarify the fundamental ideas guiding sound production and linguistic distinguishability among languages by investigating the acoustic characteristics and articulatory movements connected to each phonetic component. This study also looks into automating speech sound classification through the use of computational techniques like machine learning algorithms. We want to create reliable models that can precisely recognise and classify sounds using annotated datasets and cutting-edge signal processing techniques.**

I.INTRODUCTION

One of the many topics that contribute to language studies is the classification of sounds according to their sound characteristics. The study of the physical properties of speech is called phonetics and provides a systematic method for classifying and understanding the various sounds that make up human speech. Linguists, linguists and linguists try to unravel the complexity of speech by analyzing the quality of speech, which can give insight into the process of sound perception, production and the difference between words and words. The features that structure speech are called phonetic features, and they provide us with a way to classify and distinguish the various sounds that make up human speech. These features include many things, such as how and where they move. Our study's main objective is to investigate the application of distinctive features for automated voice recognition. Linguists utilize distinguishing features, which are a collection of attributes, to categorize.Consolidate phonemes [1,13]. To be more specific, a feature is a minimal unit that separates two maximally-close phonemes; in this case, the feature [voicE] distinguishes e x a m p l e / b / a n d / p / a r e.In comparison to the number of shared features, sounds are more frequently confused. It is thought that 15–20 separate features are enough to account for phonemes in all world languages. We will provide the results of our investigation into the three questions we previously addressed in this paper. We will first provide the results of our comparative analysis of signal representations. In light of these findings,After that, we will go over our findings and trials with acoustic attribute extraction and the application of distinguishing characteristics.Lastly, we will talk about the ramifications and offer some preliminary findings.

II.CLASSIFICATION OF SOUNDS 1.PLACE OF ARTICULATION

Indeed! The term "place of articulation" describes the location in the vocal tract where a particular speech sound is produced by restricting or obstructing airflow. One of the main phonetic characteristics used to categorize consonant sounds is this one. An outline of point of articulation and its catego1.

1.Bilabial: There is constriction between the two lips.Case in point: [p], [b], [m]

2. Labiodental: The space between the top teeth and lower lip is constricted.For instance, [f], [v]

3. Dental/Interdental: The tongue can get constricted between its higher or lower teeth.For instance, [θ] (theta) and [ð] (eth)

4. Alveolar: The alveolar ridge, or the bony ridge beneath the upper front teeth, and the tongue are constricted.For instance, [t], [d], [s], [z], [n], and [l]

5. Post-Alveolar (or Palato-Alveolar): The alveolar ridge is slightly behind where constriction takes place.For instance, [ʃ] (sh), [ʃ] (zh)

6. Retroflex: The tongue tip curls backward toward the roof of the mouth to cause constriction.Instances: [ʻ] (approximant for retroflex)

7. Palatal: The tongue and the hard palate are constricted.Instances include [j] (y in "yes") and [ʉ] (Spanish "ñ" as in "mañana")

8. Velar: The soft palate (velum) and the tongue are constricted.For instance, [k], [g], and [ʊ̈] (as in "sing")

9. Uvular: The uvula and the rear of the tongue are constricted.Examples are [ρ], [q] in Arabic. (The "r" sound in French)

10. Pharyngeal: The pharynx, or upper throat, is the site of constriction.Examples include the voiced pharyngeal fricative [ʅ].

11. Glottal: The glottis, or vocal folds, is the site of constriction.- Glottal stops: [h], [ʕ] are two examples.

2.MANNER OF ARTICULATION

The term "manner of articulation" describes how spoken sounds are produced in the vocal tract by modifying or obstructing airflow. It's yet another fundamental phonetic characteristic that's used to1.

1.Stops (Plosives): A sudden release of air after the vocal tract has completely closed.

Instances include [p], [b], [t], [d], [k], and [g].

2. Fricatives: A partial closure of the vocal tract that disrupts the flow of air.

[f], [v], [s], [z], [ʃ] (sh), [ʒ] (zh) are other examples.

3. Affricates: A pause and a fricative combined with a brief moment of total closure.

Examples are [dʒ] (j as in "judge") and [tʃ] (ch as in "church").

4. Nasals: Airflow via the nasal cavity is permitted by the full closure of the oral tract and the lowering of the soft palate (velum).

For instance, [m], [n], and [ŋ] (ng as in "sing")

5. Approximants: A somewhat smooth airflow is produced by the vocal tract narrowing without fully closing.

Examples include [w] (w), [j] (y as in "yes"), and [ɹ] (r). categorize consonants. An outline of articulation style and its categories is provided below:

8. Lateral Approximants: Airflow along the sides of the tongue due to partial closure of the vocal tract.

[l] (lateral approximant) is one example.

9. Clicks: A sequence of sounds made by the tongue and velum working together to form a suction mechanism.

For instance, [ʔ] (postalveolar click)

10. Implosives: Inward-directed airflow created by lowering the glottis.

Examples of voiced implosives are [˓].

3.VOICING

"Voicing" is another crucial phonetic characteristic that uses the vocal cords' vibration or lack thereof during sound generation to differentiate between consonants. It is important for the categorization and explanation of speech sounds. An outline of voice and its categories is provided below:

1. Spoken Sounds:noises made by the vocal cords vibrating.

[b], [d], [g], [v], [z], [ʒ], [m], [n], [ŏ], [l], [r] are a few examples.

2. Silent Sounds:

Sounds generated while the voice cords are not vibrating.

[p], [t], [k], [f], [θ], [ʃ], [s], [ʧ], [h] are a few examples.

3. Time of Voice Onset (VOT):

Voice onset time, or the interval between the start of voicing and the release of a consonant, is another way to describe voicing.

Positive VOT: After the consonant is released, voice is produced.

Negative VOT: Speaking starts prior to the consonant being released.

Zero VOT: The release of the consonant coincides with the start of voice.

III. THEORETICAL RESEARCH

Acoustic analysis, aesthetic modeling, and p sychological testing are just some of the ma ny methods used in sound classification rese arch because sound is not good. Acoustic res earch uses spectral analysis and acoustic me asurements to study the acoustic correlates o f speech, the place of speech and communic ation, and to teach music as distinct from the sounds in and within words.

Articulation Research Modeling studies the articulatory movements associated with various aspects of speech and the physiological processes involved in the production of speech sounds. Researchers can use techniques such as electromagnetic articulation and ultrasound imaging to track the movements of articulators such as the tongue, lips, and voice during speech production. This allows researchers to understand articulation coordination and differences in sound distribution. > These studies demonstrate the interaction between top-down and bottom-up processes in speech perception, revealing the importance of hearing, stimuli, and speech points in sound distribution.

1. COMPUTATIONAL METHODS

Recent advances in computing and machine learning enable new ways to classify noise by noise. Using machine learning techniques such as support vector machines, hidden Markov models and deep neural networks, situation results are processed on measured data in tasks such as speaker teaching, phone segmentation and automatic speech recognition.This computer process extracts speech data from speech symbols and uses large-scale anatomical descriptions and best practice algorithms to classify sounds into many groups. By training models on large amounts of data representing different languages and languages, researchers can develop powerful algorithms that can improve speech recognition and distinguish similar sounds.

Methods:A variety of methods have been used to identify voices by their sound, including traditional speech analysis and modern computational tools. This section provides an overview of some important techniques for learning speech-based sound classification.

2. ACOUSTIC ANALYSIS

In this process, physical properties of speech such as amplitude, frequency and time a are measured and evaluated. One application is to use techniques such as analysis, waveform analysis, and spectral analysis to find and extract acoustic cues that identify speech. Spectrograms show the frequency content of a speech signal over time and are an important tool for understanding changes in the properties of different languages.

3. ARTICULATION MODELING

Pronunciation modeling is a field of research that aims to explain the articulation movements used in speech and their connection with speech quality. Researchers can use methods such as magnetic resonance imaging (MRI), ultrasound imaging, and electromagnetic articulation (EMA) to track the movement of articulators (such as the tongue, lips, and jaw) during time articulation. Researchers can analyze articulation data to examine the coordination and diversity of articulatory movements across speakers and groups of speakers.

4.PSYCHOLINGUISTIC EXPERIMENTS

These studies investigate how listeners categorize speech and how they perceive the quality of their speech. A variety of tests are used to assess listeners' ability to determine the importance of listening and knowing the difference between many languages, including discrimination, recognition, and perceptual functions. These studies shed light on the psychological processes that lead to accurate categorization and the influence of context on perceptual categorization.

5. CALCULATION METHOD

Using machine learning algorithms and signal processing, voice counting classifies words into speech. Supervised learning algorithms (support vector machine (SVM), hidden Markov model (HMM) and deep neural networks (DNN)) are used on the recorded data to analyze the features of speech and classify sounds into predefined groups. . It uses video decoding techniques, including predictive polynomial (PLP) features and Mel Frequency Cepstrum Coefficients (MFCC) to represent the speech signal in a format suitable for distribution.

6. CORPUS-BASED ANALYSIS

This model examines different phonemes and sound patterns in different languages and languages using a general description. Use corpus linguistics techniques such as frequency analysis, cluster analysis, and correspondence analysis to identify features of speech and their distribution in linguistic material. Researchers can examine corpora to examine connections between speech and context features, including word frequency, syntactic content, and contrast.By combining various techniques, researchers can better understand speech-based sound classification and explain the auditory, articulatory, perceptual, and computational aspects of this complex phenomenon. Using these techniques, researchers can investigate differences between speakers, languages, and languages. They can also develop reliable models for phone classification, automatic speech recognition, and other speech and language technologies.

IV.RESULTS AND DISCUSSION

Classification of sounds by noise has led to many results illustrating the structure of human speech, processes related to hearing and speech, and the development of computer models for automatic speech recognition. In this section, the results and implications of research in this field are presented and discussed.

1. SPEECH ANALYSIS

This system provides a comprehensive understanding of speech differences within and between words by revealing the motor and social dynamics of acoustic interference in various speech. For example, studies have shown that spectral peaks and troughs serve as acoustic indicators for place of articulation, while formant changes provide information about articulator strength during consonant production. Additionally, changes in duration and duration of phonation begin to define different sounds, and the structure of the body and the energy of speech sounds indicate differences in the articulation process.

2. LANGUAGE DIFFERENCES

Study of different sounds. Studies comparing phone products, phone models, and phone numbers in different languages have found similarities and differences. For example, although sound difference in many languages is phonemic, language

specific phonological rules and constraints can influence the auditory cues and articulatory movements used to express vocalizations.

3. PERCEPTUAL CLASSIFICATION

The study on this topic aims to explain the role of top-down cognitive processes and bottom-up sensory processes in listeners' perception of speech and classification according to speech characteristics. It has proven effective in clinical studies reporting that listeners have faster reaction times and can more clearly distinguish between different sounds that have meaning in their language. Moreover, as perceptual research has shown, listeners' categorization of speech is influenced by the nature of the context, grammar, and general knowledge of the speech.

4. COMPUTATIONAL MODELING

Computational models of speech-based voice classification can be solved for automatic speech recognition, speaker identification, and other speech technologies. Results obtained from studies on machines show the effectiveness of supervised learning algorithms such as support vector machines and deep neural networks on speech accuracy. Through comprehensive description and demonstration of data optimization, researchers trained the model to achieve state-of-the-art performance of the job, sometimes even beyond human performance.

5. THE IMPACT OF COGNITIVE SCIENCE AND TECHNOLOGY

The scientific results and findings of sound classification according to speech characteristics have a broad impact on science, technology and grammar. They enable us to understand the universal patterns of speech and language that shape human thought and production. They also provide guidance for the development of computational models for language acquisition, speech, and communication technologies. This model may find applications in language translation, automatic speech recognition, and assistive communication devices.

In conclusion, findings and discussions in this field of research show how different sounds can be modified by computer science, psychology, phonetics, and other fields to improve communication technology and deepen our understanding of humans. words. By combining theoretical ideas with empirical data and computational methods, researchers can continue to push the boundaries of knowledge in the discipline and develop new solutions to problems in speech and language.

V.CONCLUSION

In summary, classifying sounds by sound is a complex and diverse field of research chat spans the boundaries of computer science, psychology, linguistics, and engineering. Through theoretical analysis, theoretical research, and computational modeling, researchers have made great progress in understanding the concepts behind sound classification and its impact on performance, language, communication technology, and science.Research in this field emphasizes the importance of voice quality in explaining human speech patterns and the many sounds found in different languages and languages. By analyzing acoustic correlates, joint movements, and perceptual cues associated with various speech patterns, researchers have uncovered universal and language-specific patterns of speech differences and principles that govern speech and comprehension.Additionally, thanks to advances in automatic voice segmentation models, effective speech technologies such as speaker identification, speech translation and automatic speech recognition have become possible. Through the use of machine learning algorithms and signal processing techniques, these computational methods accurately classify speech, allowing the development of better and more effective communication methods.

VI.REFERENCES

[1] Ahmed M. Abdelatty Ali*, Jan Van der Spiegel * and Paul Mueller♣*Department of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104-6390,USA,♣Corticon, Inc., 155 Hughes Rd, King of Prussia, PA 1940, USA.

[2] Andrew K. Halberstadt and James R. GlassSpoken Language Systems GroupLaboratory for Computer ScienceMassachusetts Institute of Technology Cambridge, Massachusetts 02139 USA fdrew, jrgg@sls.lcs.mit.edu

[3] Helen M. Meng, Victor W. Zue, and Hong C. Leungof Spoken Language Systems Group La boratory for Computer Science Massachusetts Institute of Technology Cambridge , Massachuset t s02139

[4] J. Ye, R. J. Povinelli, and M. T. Johnson, "Phoneme classification using naive bayes classifier in reconstructed phase space," in Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the

2nd Signal Processing Education Workshop., 2002: IEEE, pp. 37-40.

[5] T. J. Reynolds and C. A. Antoniou, "Experiments in speech recognition using a modular MLP architecture for acoustic modelling," Information Sciences, vol. 156, no. 1-2, pp. 39-54, 2003.

[6] P. Scanlon, D. P. Ellis, and R. B. Reilly, "Using broad phonetic group experts for improved speech recognition,"IEEE transactions on audio, speech, and language processing, vol. 15, no. 3, pp. 803-812, 2007.

[7] H. Kekre, A. A. Athawale, and G. Sharma, "Speech recognition using vector quantization," in Proceedings of the International Conference & Workshop on Emerging Trends in Technology, 2011, pp. 400-403.

[8] G. Kiss, D. Sztahó, and K. Vicsi, "Language independent automatic speech segmentation into phoneme-like units on the base of acoustic distinctive features," in 2013 IEEE 4th international conference on cognitive infocommunications (CogInfoCom), 2013: IEEE, pp. 579-582.

[9] G. Deekshitha and L. Mary, "Broad phoneme classification using signal based features," International Journal on Soft Computing, vol. 5, no. 3, p. 1, 2014.

[10] A. Chittora and H. A. Patil, "Classification of phonemes using modulation spectrogram based features for Gujarati language," in 2014 International Conference on Asian Language Processing (IALP), 2014: IEEE, pp. 46-49.

[11] H. H. Nasereddin and A. A. R. Omari, "Classification techniques for automatic speech recognition (ASR) algorithms used with real time speech translation," in 2017 Computing Conference, 2017: IEEE, pp. 200-207.

[12] S. Salim, G. Deekshitha, A. George, and L. Mary, "Automatic Spotting of Vowels, Nasals and Approximants from Speech Signals," in 2018 International CET Conference on Control, Communication, and Computing (IC4),2018: IEEE, pp. 272-277.

[13] A. V. Anand, P. S. Devi, J. Stephen, and V. Bhadran, "Malayalam Speech Recognition system and its application for visually impaired people," in 2012 Annual IEEE India Conference (INDICON), 2012: IEEE, pp. 619-624.

[14] G. Deekshitha, K. H. Hathoon, and L. Mary, "A Novel Two-Stage System for Spotting Fricative and Plosive Regions from Continuous Speech," in 2018 International Conference on Communication and Signal Processing (ICCSP), 2018: IEEE, pp. 0760-0764.