

## Machine Learning for Speech and Voice Analysis in Aphasia: Current Trends and Challenges

*Sonawane Khushbu<sup>1</sup> and Dr. Shripal Shah<sup>2</sup>*

*<sup>1</sup> Atmanand Saraswati Science College*

*<sup>2</sup>Shree Ramkrishna Institute of Computer Education and Applied Sciences*

### Abstract

This communication disorder known as aphasia and it normally happen when human has a stroke or brain damage in which human has difficulty to speak, understand language and communicate clearly. This topic has enhanced in computer-based language and speech research. Traditional diagnostic evaluations are clinically trustworthy; however, they depend heavily on expert clinically informed interpretation, take a lot of time and must be done by trained professionals. As a result, they may delay the start of therapy. Luckily, Machine Learning (ML) is currently providing us with new means of automatically browsing speech and voice data, which may facilitate and objectify the process of evaluation.

This study included analysis of seven studies that apply various tricks of Machine Learning to identify and treat aphasia. The study discusses not only combination learning methods, but it incorporates NLP based severity prediction, automated diagnosis, deep learning-based speech analysis models, and hybrid learning approaches. This study indicates that the use of machine learning can support correct diagnosis, On-time treatment and patient specific treatment planning. yet, we face issue like a limited dataset and the risk of overfitting still remains. This paper suggest that machine learning has the potential to grow healthcare services in the future and become accessible to a large population. It can also help people with different languages Hence diagnosis and treat patient are more effective than using traditional clinical methods.

**Keywords:** Aphasia; Machine Learning; Speech and Language Processing; Natural Language Processing; Deep Learning.

### 1. Introduction:

Aphasia is a language disorder, which is normally seen in people, who have had a stroke and brain injury. In this condition, patient can face problem like they can't understand language, can't generate speech, and can't express their feelings. This condition described in IEEE (2021) [9]. Additionally, an important study by Tóth et al. (2021) [8] is abnormal speech pattern, which may consider as major issue that affected on patient daily life because they can't proper communicate with others, and it can also impact on person's personality. In real life doctors are use basic method like Western Aphasia Battery (WAB) and Boston Diagnostic Aphasia Examination (BDAAE). Basically, these types of tests are widely used, they are lengthy and need expert person to perform properly as noted by Day et al. (2021) [3].

Now days, using of advance Machine Learning (ML) it easy to automatically identify aphasia. ML can help to identify patient intensity, can analyse their speech and also can identify types of aphasia Maddipatla (2022) [6], showed that merging different models together and make diagnosis more accurate, while Nambiar et al. (2020) [7] showed the importance of deep learning, Ashwini et al. (2023) [1] showed that ML based system can be helpful in personalized treatment for patients and their progress.

In this review include studies from 2020 onwards and focused on the use of ML, it's advantages, effectiveness and remaining challenges.

## 2. Literature Review:

I have reviewed several research papers, and in that paper, I read that ML is very helpful for aphasia patients to identify their types and how severe they are. In this paper **IEEE (2021) [9]**. explain different types of ML models, like supervised, unsupervised, and reinforcement learning, which are helpful for aphasia patients. Used ML models are very useful idea but they have not yet been tested on any real patient, so basically these is the only idea and that models are use in real-life, they discussed only in their paper.

**Maddipatla (2022) [6]** performed various practical activities and developed system called as Class Aphasia. In this system included multiple ML models, that identify the accuracy of aphasia and the severity of the aphasia. The system works with more than 25,000 speech samples and gives high accuracy as a result, but they didn't use this system on real patients, so the result may not apply to everyone, and the model might be overfit.

**Day et al. (2021) [3]** studied how sentences are made, what is meaning of the word, and how many times words are used. They integrate these facts with the machine learning algorithms such as Random Forests and Neural Networks. Using these models, they assumed how serious the person's aphasia is. They found one problem that the model gives perfect accuracy on normal patient data, but does not give perfect accuracy on severe aphasia patients data.

**Khan et al. (2017) [5]** work on a system called You Speak, We Detect. In this system using methods like speech recognition and signal processing are used to identify the types of aphasia, like Wernicke's and anomic aphasia. They study how people speak, how they find out that word, while speaking. This system worked well, but there is a shortage of data, and it only predicts two types of aphasia.

**Jothi et al. (2020) [4]** The Reviewer gave idea that therapy should be personal for each and every aphasia patient and also using three types of learning methods: supervised, unsupervised, and reinforcement. Experiments on aphasia patients never been done yet, these are only suggestions.

**Zhong (2024) [10]** AI/ML are very useful for aphasia patients. According to a study, these tools have many advantages like, recently take feedback, benefit of many users, and give treatment on data of patients. They also face many challenges, like there is no support for different languages, they have no experiment on clinical data, and not detect emotions.

Recently, **Ashwini (2023) [1]** team tried deep learning and find out language problems, like apraxia, and aphasia. Their main motive is using Neural Network instead of old methods like GMM, HMM and they used CNN, RNN and also mix models and one thing that Neural Network methods are learn from the data.

The technology is very developed and progressive but there are many issues, like it's consumed much computing power, these approaches are time-consuming and have not been tested and lack experiments on real patients. So, that's why we are not sure how it will be work on actual data.

*Table 1: Review of previous research On Aphasia*

Study (year)	Methods	Dataset (summary)	Key Findings / Results	Identified Limitations
<b>IEEE Survey (2021) [9]</b>	Supervised, unsupervised, and reinforcement-learning models	Invalid	ML used for study speech problems	Only theory, no testing
<b>Maddipatla (2022) [6]</b>	Ensemble model (Class Aphasia)	Maybe 19,425 speech samples, 6,474 for training	high accuracy in some study	overfitting; no clinically tested
<b>Day et al. (2021) [3]</b>	NLP features: Random Forest and Neural Network	Theoretical dataset	73% accuracy report	Unbalanced data and few severe case
<b>Khan et al. (2017) [5]</b>	ASR + DSP	Small patient dataset	Detect Wernicke's and Anomic only 2 types of aphasia	Small sample, limited subtypes
<b>Jothi et al. (2020) [4]</b>	Hybrid ML concept	Not needed	personalized therapy for each patient	Just concept, not practically work and any proof
<b>Zhong (2024) [10]</b>	AI/ML	Doesn't apply	real-time and large data can help	Missing emotional and language factors; no any clinical trials
<b>Ashwini et al. (2023) [1]</b>	Deep-learning (CNN, RNN, hybrid) models	Many ASR dataset and clinical studies datasets	Shift to Deep Learning models	Need heavy computing and not tested in clinic

### 3. Objectives:

The main Objective of this Review paper is to understand how Machine Learning, Deep Learning, are used to detect, study and treat aphasia. It focuses on many different methods such as speech and processing, NLP-based analysis, ensemble learning, and deep learning models that recent researchers have used in papers.

In this paper also tries to compare traditional clinical test with ML based systems. Another goal is pointing the strength or weakness of models, approaches and main problem is that

small dataset, lack of clinical testing. At last, this review looks how artificial intelligence can help in more personalized therapy to aphasia patients.

#### **4. Research Methodology:**

We use Machine Learning to test and study aphasia. we can't use raw audio data, so that's why we convert audio data into numbers that computers can understand easily. After that, researchers can study how the voice sounds, how much time it takes while speaking, and how many mistakes can be made while speaking. Nambiar et al. (2020) [7] briefly explain this concept; it's similar to what we learn in the classroom: before training the model, we take audio data and extract useful features.

Finally, we collect all those features and remove those that are not useful. Then, the useful features are used in SVMs, Random Forest models, or Deep Learning networks for predicting severe aphasia or what pattern follows for predicting aphasia. Day et al. (2021) [3] find that when language features and ML can combine, we get better predictions.

Deep learning is a recent study; their system can easily understand speech without a lot of manual audio processing Ashwini et al., (2023) [1]. Some researchers, like Maddipatla (2022) [6] shown that merging multiple models can predict audio more accurately and dependably.

Some systems are not just shown aphasia level but also help with treatment. The system gave exercises depending on the patient's improvement level. Tasks are more difficult when patients perform well and more straightforward when patients perform poorly. Jothi et al. (2020) [4] explained that merging different learning methods can help make therapy more personalised for each patient.

#### **5. Research Problem:**

Machine learning can find and treat aphasia more efficiently, but it also have some problems. For example, Day et al. (2021) [3] used NLP and Machine Learning to show good result, but that system works only on mild or moderate aphasia. This means that their system does not work properly for persons with severe aphasia.

In the exact way, Maddipatla (2022) [6], using the Class Aphasia model, got nearly perfect accuracy, but the model was not tested by other researchers and was not tested in a real hospital. So, there is a chance that the model doesn't work properly in a real-life situation, and it may suffer problems like overfitting.

Jothi et al. (2020) [4] suggested that therapy should be personalized, using multiple types of ML algorithms together, but these are only a theoretical purpose. Even, they did not try in a real situation. Also, in another study, the You Speak, We Detect system Khan et al., (2017), they were able to find types of aphasia, but they found only two types of aphasia, not more than that. This means the system needs to be more testing, especially when working with multiple patients from different language background.

I read these papers, and I find the main major issues that all models and systems do not have proper clinical testing, also Zhong (2024) [10] and Ashwini et al. (2023) [1] in this paper talking about how widely use Machine Learning for aphasia patients, but the major issues is don't apply this model on real-life patients.

In the end, most problems are that we don't have open and varied datasets. Mostly, researchers use a small dataset or only a one-language dataset. So, that's why Machine Learning models cannot work properly for all patients.

We must develop a large data set of many languages and every type of aphasia in the future. We also desire to put these models to test with actual patients in actual conditions.

If we want Machine Learning tools to be useful for diagnosis and treatment in hospitals, then these steps are very important.

## 6. Findings:

When, we study all papers together, it becomes clearer that machine learning has lots of power to improve aphasia and it is treated Ashwini et al. (2023) [1]; Day et al. (2021) [3]. Researchers have shown that Machine Learning, Deep Learning, Neural Network, SVM, and Random Forest can understand speech and language patters easily Ashwini et al. (2023) [1]; Maddipatla (2022) [6]; Nambiar et al. (2020) [7]. Also, these systems are able to detect aphasia and estimate how serious it is much faster than traditional Day et al. (2021) [3]; IEEE (2021) [9]. Many papers also show that when language features are combined with ML so, result is more accurate Day et al. (2021) [3]; Maddipatla (2022) [6]. Using more than one model at the same time it's give batter output Maddipatla (2022) [6]. DL can more helpful because it's learned directly from the speech data on its own, it does not depend much on hand made features Ashwini et al. (2023) [1]; Nambiar et al. (2020) [7].

These study shows some real-world problems. Most of the research on small dataset and only one language Maddipatla (2022) [6]; Zhong (2024) [10]. Some systems are not tested on real patients in hospital Jothi et al. (2020) [4]; Maddipatla (2022) [6]; Zhong (2024) [10]. Some model gives better accuracy on mild case but in severe case model face many issues Day et al. (2021) [3]. Basically, these is risky because model can work properly only on trained dataset, but not work properly on real situation Maddipatla (2022) [6]; Zhong (2024) [10].

## 7. Conclusion:

Nowadays, Machine Learning can be helpful for understand aphasia speech issues. Many models added simple machine learning models to more powerful Deep Learning system. In most of the cases, models can identify the condition and give treatment ideas.

There are still several challenges to be addressed, as many studies use small dataset, especially when patient have severe aphasia. The models do not work for all people, as the results are different when the languages are different. Major issue is that models have not been tested in hospital or large clinical data.

Right now, Machine Learning can't fully replace with traditional test, but it may speed up and improve the process for doctors.

Say it simple words, Machine Learning gives us strong base for understanding speech disorder, however more research is needed before it can be used in real-life diagnosis and therapy.

## 8. Suggestions:

From this review, it is clear that future research needs to move in more practical direction. Researchers should work on big and large dataset in which include multiple languages and multiple types of aphasia cases. It is also important that Machine Learning system are tested in hospitals with real patients. Thus, we get to know the performance of the system on actual patients. Also, mainly focus on severe aphasia patient because it is missing in many studies.

Also, hybrid and ensemble methods for get more accuracy and stability. And new system also able to work with multiple languages and pay attention to emotional and thinking

related parts of speech. Very important computers experts and medical professional need to work closely together. So, that ML based system can truly useful and safe for real diagnosis and therapy.

## 9. References:

1. Ashwini, P., Bharathi, S. H., & Raghavendra, M. K. (2023). A survey of speech recognition deep learning methods for apraxia of speech disfluency identification application. In 2023 IEEE 5th International Conference on Advances in Electronics, Computers and Communications (ICAIECC) (pp. 1–7). IEEE. <https://doi.org/10.1109/ICAIECC59324.2023.10560190>
2. Chan, W., Jaitly, N., Le, Q. V., & Vinyals, O. (2016). Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. IEEE ICASSP, 4960–4964. <https://doi.org/10.1109/ICASSP.2016.7472621>
3. Day, M., Dey, R. K., Baucum, M., Paek, E. J., Park, H., & Khojandi, A. (2021). Predicting severity in people with aphasia: A natural language processing and machine learning approach. In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 2299–2302). IEEE. <https://doi.org/10.1109/EMBC46164.2021.9630694>
4. Jothi, K. R., Sivaraju, S. S., & Yawalkar, P. J. (2020). AI-based speech language therapy using speech quality parameters for aphasia person: A comprehensive review. In 2020 Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1263–1271). IEEE. <https://doi.org/10.1109/ICECA49313.2020.9297459>
5. Khan, M., Silva, B. N., Ahmed, S. H., Ahmad, A., Din, S., & Song, H. (2017). You speak, we detect: Quantitative diagnosis of anomic and Wernicke’s aphasia using digital signal processing techniques. In 2017 IEEE International Conference on Communications (ICC) (pp. 1–6). IEEE. <https://doi.org/10.1109/ICC.2017.7996800>
6. Maddipatla, J. (2022). ClassAphasia: An ensemble machine learning network to improve aphasia diagnosis and determine severity. In 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 2314–2317). IEEE. <https://doi.org/10.1109/BIBM55620.2022.9995616>
7. Nambiar, V., Hemanth, D. J., & Anitha, J. (2020). Deep learning for speech and language processing in healthcare: A review. Computer Methods and Programs in Biomedicine, 196, 105627. <https://doi.org/10.1016/j.cmpb.2020.105627>
8. Tóth, L., Gosztolya, G., & Grósz, T. (2021). Speech recognition for atypical speech: A survey. *Computer Speech & Language*, 72, 101298. <https://doi.org/10.1016/j.csl.2021.101298>
9. IEEE (2021). Analyzing machine learning algorithms for speech impairment related issues. In 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 1183–1187). IEEE. <https://doi.org/10.1109/SPIN52536.2021>
10. Zhong, X. (2024). AI-assisted assessment and treatment of aphasia: A review. *Frontiers in Public Health*, 12, 1401240. <https://doi.org/10.3389/fpubh.2024.1401240>