Sarcouncil Journal of Multidisciplinary

ISSN(Online): 2945-3445

Volume- 05| Issue- 07| 2025

Review Article **Received:** 06-06-2025 | **Accepted:** 29-06-2025 | **Published:** 19-07-2025

Voice AI Agents: The Technical Backbone of Modern Speech Interfaces

Varghese Paul
Independent Researcher, USA

Abstract: Voice AI agents are a moment in technological development that opens doors between people and technology in the most natural form, voice. More than ever before, it is easy to bridge gaps in human communication when those barriers are based on distance, differences in language, and the ability to communicate face-to-face. This report takes a closer look at the inner wiring of voice interface systems and breaks down the three important mechanisms that make or break modern voice assistants. The adventure starts with Voice Activity Detection, a guard keeping human utterances away from the chaos of background noise, by non-stationary sound processing mechanisms. Moving deeper, the transcription engine transforms sound waves into meaningful text via intricate neural pathways and linguistic frameworks. Completing this technological symphony, speech synthesis converts digital responses into remarkably human-like vocal expressions. The fusion of these elements has sparked transformation across business communication centers, hospital record-keeping systems, power-conscious gadgets, and tools for people with disabilities. Each application domain brings unique hurdles that can be overcome through clever engineering solutions—distributed processing frameworks, specialized computing chips, and secure communication pathways. As voice technologies mature and spread, the landscape of human-machine connection continues shifting toward more natural, accessible experiences tailored to countless real-world situations where traditional interfaces fall short.

Keywords: Speech Recognition, Natural Language Processing, Voice Synthesis, Neural Vocoders, Multimodal Interfaces.

INTRODUCTION

Remember the first time a computer understood your voice? Probably frustrating, filled with endless repetition and machine confusion. Today, millions chat with Alexa, Siri, and chatbots working as customer service agents, and they understand us. This was not an instant change.

Voice technology had decades of simple issues on its hands, making out a human voice against the hash of a noisy room with different accents. It did not seem natural to talk to the machine, but it was like talking to a robot. Engineers overcame these adversaries in successive small steps instead of a eureka experience. Technology only became more useful than amusing; that is, technology passed that point of being interesting yet frustrating to a point where it became very helpful in everyday use, leading to market growth taking off like never before.

Digital signal processing breakthroughs finally tamed the background noise problem, creating systems that work reliably even with dogs barking or music playing nearby (Lee, S, 2025). Without this fundamental improvement, voice assistants would remain tethered to perfectly quiet rooms – technically impressive but practically useless. It is much more important to achieve lab perfection than real-world robustness.

This seemingly magical process is achieved with the use of three technological pillars: Voice Activity Detection determines whether or not it is actually human voice making noise; Transcription transcribes the sound waves into actual words and sentences; Text-to-Speech translates digital text back into surprisingly human-sounding words. The parts are developed using drastically different technical solutions, but are designed to work as a single whole.

Testing these systems demands holistic thinking beyond simple accuracy metrics. Response timing, naturalness, and contextual understanding – these subjective elements often determine whether users embrace or abandon voice technology. Industry evaluation frameworks increasingly incorporate these human factors alongside traditional technical measurements (Saxena, S. 2025). Numbers tell only part of the story when human perception ultimately determines success or failure.

This analysis explores the inner workings behind modern voice interfaces, revealing not just individual components but their complex interplay. The hardware-software relationship becomes particularly crucial as deployment environments diversify beyond controlled settings. Neural network architectures revolutionized every aspect of this field, replacing fragile rule-based systems with models that learn directly from speech patterns. The latest systems handle diverse accents and speaking styles without requiring individual calibration, adapting on the fly to new speakers. These calculations are more often and more decentrally performed by specialized processors, allowing for natural conversations without being continuously connected to the cloud or being in battery depletion mode. The rate of innovation keeps growing faster and faster, redefining

***Corresponding Author:** Varghese Paul

DOI- https://doi.org/10.5281/zenodo.16149167

possibilities of human-machine communication every time.

VOICE ACTIVITY DETECTION (VAD): IDENTIFYING HUMAN SPEECH

The first problem of voice processing comes in the very beginning: the separation of human speech and the ambient noise. Voice Activity Detection acts as a guardian, or rather a filter, of this process through which complex signal analysis is used to cherry-pick valuable utterances out of the ambient acoustic environment. The sound on these modern systems is examined in different ranges of frequencies, and the sound is searched for a particular signature that will help differentiate between a human voice and a car horn, music, or the drone of a machine. The magic comes in the form of hybrid solutions that combine conventional signal processing with models of machine learning trained with thousands of different examples in audio. Successful systems must juggle numerous acoustic features simultaneously – energy distribution patterns, zero-crossing rates, cepstral characteristics, and harmonic relationships – all contributing vital clues about where speech begins and ends (Beritelli, F. *et al*., 1998). The acoustic world presents an endless variety, requiring systems that adapt to shifting environments rather than relying on rigid thresholds.

Voice detection serves multiple practical functions beyond simply identifying speech. Battery life in mobile devices depends critically on efficient processing – constantly analyzing audio would drain power reserves rapidly. By activating full speech recognition only when necessary, detection systems dramatically extend operating time while maintaining responsiveness. Error prevention represents another vital benefit, as mistaken activation creates frustrating user experiences. Well-designed detection frameworks strike a delicate balance, remaining sensitive enough to catch genuine commands while rejecting false triggers from similar-sounding noises. Studies show detection systems dramatically reduce misactivation while maintaining quick response to actual speech inputs (Way With Words,). The difference between a frustrating voice assistant and a helpful one often hinges on this initial detection stage.

Adaptation to changing environments represents perhaps the most impressive aspect of modern detection systems. Walking from a quiet office into a crowded street completely transforms the acoustic landscape, yet effective voice systems continue functioning without missing a beat. This adaptability stems from neural network designs incorporating contextual awareness mechanisms that continuously shift attention toward speech-relevant features while ignoring distractions. Flexibility is of particular interest to mobile applications because users have to go through a widely varying acoustic environment during the day (Beritelli, F. *et al*., 1998). The detection thresholds in real-time, to constant background noise-dependent thresholds, are set behind the scenes by extrapolation of low-order knowledge in sophisticated algorithms that do not need manual adjustments when moving between extremely different signal-to-noise conditions.

The usefulness of advanced voice recognition is best seen in difficult acoustic conditions, such as call centers with several conversations simultaneously, cars with traffic noise and music playing, or places with a lot of competing sounds. One of the most challenging automotive applications, driver recognition is critical to automotive applications that need to parse passwords out of driver commands, passenger jabber, engine roars, and along with the entertainment system. Outside consumer use, voice detection is also fundamental to communication infrastructure, especially Voice over Internet Protocol infrastructure, where bandwidth consumption must primarily be limited to when active speech is being carried. These systems dynamically allocate network resources based on speech detection, significantly reducing data usage without compromising call quality (Way With Words,). As voice interfaces spread into increasingly diverse environments, detection technology continues evolving to handle previously impossible acoustic scenarios, gradually removing environment limitations from voice interaction.

Publisher: SARC Publisher

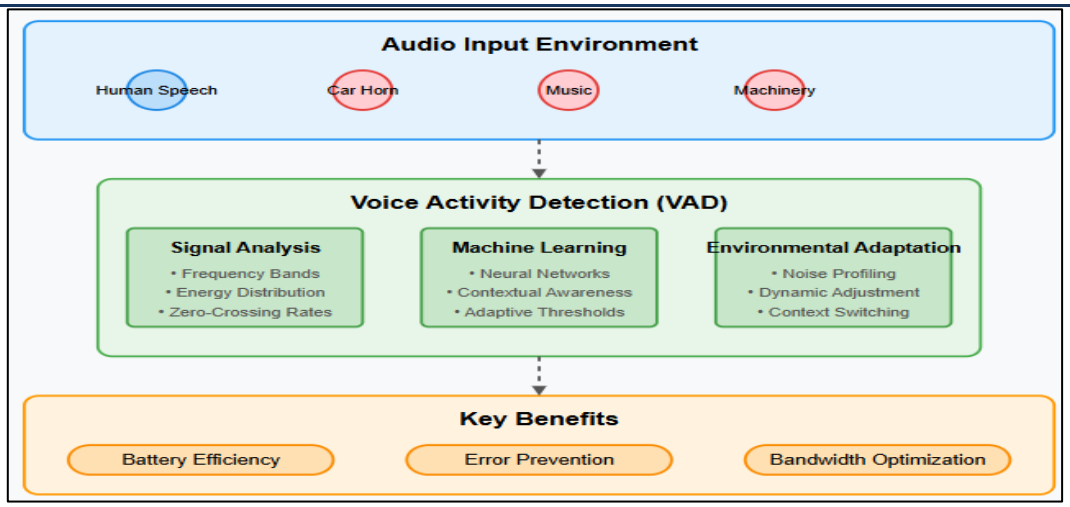

Fig 1: Voice Activity Detection (VAD) System Architecture (Beritelli, F. *et al*., 1998; Way With Words,)

TRANSCRIPTION (SPEECH-TO-TEXT): CONVERTING AUDIO TO DATA

When speech is identified, a monumental process of deriving meaning out of the sound waves starts. One of the keys to unlocking speech portals and forwarding the voice into an integrated device is the transcription system, which transforms air to vibration to organize data that can be translated. Contemporary systems of speech recognition are barely similar to their more primitive precursors, and have now been radically improved by the use of deep neural networks trained on large datasets. The three-letter words spoken by the ancestors were always a challenge to master even in controlled environments, and there was no chance of the elders knowing how to deal with real-life conversations consisting of multiple languages. But with the kind of engines in use today, eloquence is retained, and no foreign language has to be learned by heart to have a conversation in today's date. The shift toward transformer-based networks capable of tracking long-range dependencies has proven particularly transformative, allowing systems to maintain context across extended utterances. These advancements shine especially bright when processing natural conversation with its inherent messiness – false starts, hesitations, interruptions, and varying speech rates no longer derail recognition (Xiong, W. *et al*., 2016). Spoken language, after all, follows different patterns than written text, requiring specialized approaches rather than mere adaptation of text processing techniques.

Effective transcription derives its basis in acoustic modelling, which builds structures linking the patterns of sound with units of language, even though there exists variation among speakers. Indeed, there is a bit of difference in words pronounced by every person due to the accent, physiology, and the style a person speaks, but recognition systems need the ability to operate with this flavour of diversity without doing a per-person calibration. To play a significant role, feature extraction is usually performed, which involves transcoding raw audio into specific representations such as mel-frequency cepstral coefficients that accentuate particular perceptually meaningful variations and suppress redundant variation. State-of-the-art models deploy attention mechanisms allowing dynamic focus on the most informative portions of input signals, dramatically improving recognition of challenging speech patterns. This architectural approach has virtually eliminated the need for explicit speaker adaptation procedures that older systems required, allowing immediate high-quality recognition across diverse speaker populations (Jurafsky, D. and Martin, J. H. 2025). The impact extends beyond convenience to genuine inclusion, with significantly improved performance for non-standard speech patterns, including regional accents and speech affected by physical conditions.

Language modeling represents the complementary half of transcription systems, providing contextual frameworks that disambiguate similar-sounding utterances based on linguistic probability. Humans easily distinguish between "recognize speech" and "wreck a nice beach" despite acoustic similarity because context and language knowledge constrain possible interpretations. Modern language models capture these constraints through massive neural networks trained on billions of text examples, learning probabilistic relationships between words at multiple scales from local phrases to document-level themes. The marriage between acoustic

Publisher: SARC Publisher

processing and language modelling produces particularly powerful results in specialized domains where specific terminology and speech patterns dominate (Xiong, W. *et al*., 2016). Legal proceedings, medical documentation, and financial services all benefit from domain-adapted models that understand the specialized language and typical phrasing patterns unique to those fields, dramatically outperforming generic transcription when working within their intended domains.

Multilingual capability represents another remarkable achievement in modern transcription, enabling support for diverse languages without requiring completely separate systems. The diversity of languages in human cultures can be astonishing in the number of sounds, forms, and patterns that are utilized, but current domains deal with dozens of languages using similar or identical structures. This is a cross-language performance which is built through advanced transfer learning methods where the nature of existing similarities between languages is exploited and the knowledge that is obtained in the data-abundant languages is used to improve performance in the ones where limited training data is available. These approaches have dramatically improved language coverage, enabling practical deployment in environments where multiple languages might be encountered, such as international business communications or global customer service operations (Jurafsky, D. and Martin, J. H. 2025). The societal impact extends beyond convenience to genuine democratization of technology access, reducing the historical bias toward major commercial languages in speech technologies.

Performance measurement in transcription systems primarily revolves around word error rate, though this metric tells only part of the story. Applications requiring precise command interpretation, such as healthcare documentation or financial services, deploy specialized engines with additional error correction mechanisms and domain-specific language models. These systems incorporate secondary validation layers performing contextual verification through linguistic and domain-specific rule checking, identifying likely transcription errors based on semantic inconsistency rather than just acoustic confidence (Xiong, W. *et al*., 2016). Such sophisticated approaches prove particularly valuable in high-stakes environments where transcription errors could have significant consequences, misinterpreted medical instructions, or incorrect financial figures, potentially causing serious harm. Interestingly, perfect word-level transcription may not always be necessary, with emerging evaluation approaches focusing more on semantic correctness and intent recognition, acknowledging that human communication contains redundancy, allowing meaning to survive despite imperfect word capture (Jurafsky, D. and Martin, J. H. 2025). This shift reflects a maturing understanding that the ultimate goal lies in extracting meaning rather than creating perfect text renderings of speech.

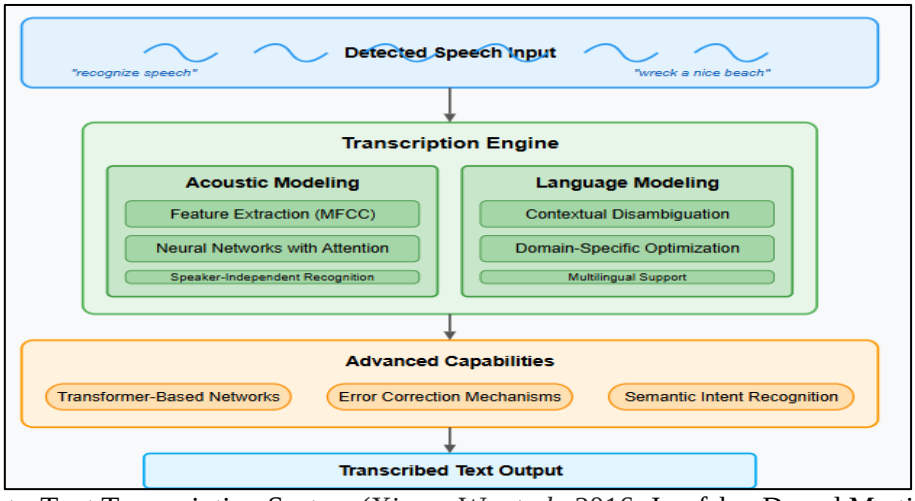

Fig 2: Speech-to-Text Transcription System (Xiong, W. *et al*., 2016; Jurafsky, D. and Martin, J. H. 2025)

TEXT-TO-SPEECH (TTS): SYNTHESIZING NATURAL VOICE OUTPUT

Do you recall the old sci-fi movie with computerized robotic voices? And that is how the early text-to-speech sounded: monotone, mechanical, and anyone could hear how it sounded fraudulent. Then, skip to modern times, and digital voices are very eerily human-sounding. Such outstanding costs were not achieved gradually but as a result of a series of technological revolutions. The earliest commercial systems worked like

Publisher: SARC Publisher

verbal Frankenstein monsters, chopping pre-recorded human speech into tiny fragments and then stitching them back together to form new words. This "concatenative" approach produced somewhat understandable speech but sounded choppy and unnatural at sentence boundaries. The voice quality fell apart completely when speaking anything not anticipated by the original recordings.

Everything changed when neural networks entered the picture. Modern systems scrapped the cut-and-paste approach entirely, instead learning the fundamental patterns connecting written language to spoken sounds (Ren, Y. *et al*., 2020). It is more like the contrast between learning certain phrases of a foreign language, versus learning the grammar and the rules of pronunciation of that language. Neural TTS does not simply play back human recordings; it synthesizes speech out of nothing using patterns it has picked up.

The secret sauce behind truly natural-sounding speech lies in neural vocoders - specialized components that generate the actual waveforms making up speech sounds. These mathematical models capture those subtle variations that make human speech sound alive: tiny pitch fluctuations, microscopic timing differences, and natural transitions between sounds. Different technical approaches emerged, each with its own strengths - autoregressive models produced beautiful audio but ran slowly, while GAN-based models sacrificed some quality for speed. Engineers recently managed to shrink these sophisticated models to run on phones and smart speakers without constant internet connections (Zen, H. *et al*., 2013). That breakthrough explains why your virtual assistant doesn't sound so robotic anymore, even when WiFi goes down.

Beyond just sounding human, modern speech systems capture emotional qualities that completely transform the listening experience. Want your navigation app to sound excited about your destination? Or your virtual assistant to express sympathy when you're having a rough day? Emotional synthesis adjusts dozens of subtle voice parameters - pitch range, speaking rate, breath patterns - mimicking how human emotions physically change our voices. Advanced systems model actual changes in virtual vocal tract shapes, simulating how emotions physically alter human speech production (Ren, Y. *et al*., 2020). The difference between flat, robotic delivery and emotionally appropriate speech often determines whether an interaction feels mechanical or natural.

Voice customization might be the most transformative recent development. Creating a new synthetic voice once required months of studio recording and processing. Modern systems need just minutes of sample audio to capture someone's unique voice characteristics. Transfer learning techniques - basically teaching new voices by leveraging knowledge from previous ones - made this dramatic improvement possible (Zen, H. *et al*., 2013). This democratization matters tremendously for accessibility applications, letting people who've lost their ability to speak create personalized synthetic voices matching their identity rather than using generic alternatives. The psychological comfort of maintaining one's vocal identity cannot be overstated.

Measuring voice quality presents unique challenges, since technical perfection doesn't necessarily correlate with human preference. The industry relies heavily on Mean Opinion Score tests, where human listeners rate voices on naturalness, clarity, and appropriateness (Ren, Y. *et al*., 2020). These subjective tests catch issues that pure signal analysis misses - voices might be technically perfect, but somehow unsettling (the famous "uncanny valley" problem). Researchers recently developed automated metrics that correlate well with human judgments, speeding up development cycles without sacrificing quality (Zen, H. *et al*., 2013). Different applications prioritize different voice qualities - navigation systems value clarity above all, while entertainment applications might prioritize expressiveness over perfect pronunciation. There's no single "best" synthetic voice; there are just different voices optimized for different purposes.

Table 1: Evolution of Text-to-Speech Technology: Quality vs. Customization Time (Ren, Y. *et al*., 2020; Zen, H. *et al*., 2013)

TTS Generation	Technology	Quality Score (1-10)	Emotional Range	Voice Customization Time	Offline Capability	Application Priority
Early Systems	Robotic/Mechanical	2	None	N/A	Yes	Basic Communication

Publisher: SARC Publisher

Concatenative	Pre-recorded Fragments	4	Limited	Months	Yes	Word Accuracy
Neural - Early	Basic Neural Networks	6	Basic	Weeks	Limited	Natural Rhythm
Neural - Advanced	Neural Vocoders	8	Multi-emotion	Days	Yes	Human-like Quality
Current Gen	GAN/Transformer-based	9	Full Spectrum	Minutes	Yes	Context Appropriate

TECHNICAL APPLICATIONS AND IMPLEMENTATION CONTEXTS

Voice technology has gone viral and has changed the way people live, how businesses are conducted, and how people can access services. Customer service has seen perhaps the most dramatic real-world impact. Gone are the endless "press 1 for sales" phone menus, replaced by systems that understand what customers need. These aren't simple chatbots - they're sophisticated systems that remember previous conversations, know your account details, and solve problems without human involvement. Banks, phone companies, and retailers have jumped on this technology, handling endless routine questions through natural conversations instead of forcing customers through frustrating button-mashing exercises (Mathur, A. *et al*., 2017). Beyond just saving money, these systems work 24/7, maintain consistent quality during peak times, and solve problems immediately rather than forcing callbacks. The difference between clunky automated systems and smooth conversational interfaces dramatically impacts whether customers walk away satisfied or frustrated.

Doctors and nurses have embraced voice technology for an entirely different reason - escaping mountains of paperwork. The medical wards have their problems: the unique vocabulary (read aloud the word pneumonoultramicroscopicsilicovolcanoconiosis and remember how it is pronounced correctly) and speaking patterns. The modern systems transcribe what doctors say, but also order the resulting data into correct medical records, automatically populating the correct data fields with symptoms, diagnosis, and therapy plans. Patient care has improved through voice assistants that remind about medications, explain treatment instructions, and answer basic health questions through conversation rather than easily ignored text messages. Research consistently shows patients, especially older folks and those who struggle with technology, respond better to voice reminders than written ones (World Health Organization, & United Nations, 2022). Building these healthcare systems requires extreme attention to privacy laws and security needs. One data leak could spell disaster, forcing developers to balance robust security with systems that still work efficiently and feel natural to use.

Making voice technology work on small devices presents completely different headaches. Limited battery life, tiny processors, and minimal memory force brutal trade-offs between performance and practicality. Engineers employ clever tricks - simplifying neural networks while preserving accuracy, reducing mathematical precision requirements, and streamlining processing pathways. These techniques squeeze voice capabilities into devices that couldn't possibly run full-sized systems. Cars present perhaps the toughest environment - engine roar, road noise, wind, music, and passengers talking simultaneously create acoustic nightmares for voice recognition (Mathur, A. *et al*., 2017). Advanced microphone systems with sound-shaping programs are used, which can isolate the driver's speech signal out of this noise and accurately identify commands given at autobahn speeds with the radio on full blast. When voice features can be occasionally utilized, they are dismissed as gimmicks; when voice features can always be utilized, they are daily necessities.

Accessibility applications are the brightest example of how voice technology can change. Voice control gives independence to individuals in the physically challenged group to operate computers, control home environments, and access content without physical manipulation to operate controls, which are dependent on fine motor skills. The systems are useful in helping the speech disabled because they have been optimized to recognize various speech patterns, such as patterns in a cerebral palsy or ALS patient. Voice interfaces can be very important to visually impaired individuals as they give them access to information such as screen reading, environmental description, and navigation support. Real-world studies consistently show that properly designed voice tools dramatically improve quality of life for people with disabilities, enabling participation in

Publisher: SARC Publisher

activities otherwise impossible (World Health Organization, & United Nations, 2022). Creating truly inclusive voice applications requires a genuine commitment to universal design principles and extensive testing with actual users having diverse needs and abilities, not just checking boxes for compliance but ensuring genuine usability across different situations and conditions.

Each application environment brings distinct technical hurdles requiring creative engineering solutions. Edge computing represents one popular approach, splitting processing between devices and cloud servers based on specific needs. Time-sensitive operations happen locally for immediate response, while more intensive processing occurs in powerful remote data centers. Another key benefit is that entirely custom voice chips are optimized to perform the rather complicated matrix math at the core of neural networks (Mathur, A. *et al*., 2017). The specialized new processors deliver vastly improved results with much less power than the general-purpose computer chip, allowing advanced speech processing with no battery draining. Data protection presents another critical challenge, with encryption systems safeguarding voice information during transmission. Specialized security techniques balance protection needs against response time requirements - excessive encryption delays would make conversation impossible, while inadequate security risks exposing sensitive information. As voice interfaces continue expanding into new territories, each field requires unique adaptation addressing specific constraints and requirements, gradually replacing traditional interfaces in situations where keyboards and screens prove impractical or inaccessible (World Health Organization, & United Nations, 2022).

Table 2: Voice AI Applications across Industries: Benefits vs. Technical Challenges (Mathur, A. *et al*., 2017; World Health Organization, & United Nations, 2022)

Sector	Primary Use Case	Key Benefit	Technical Challenge	Implementation Approach
Customer Service	Automated Inquiries	24/7 Availability	Contextual Memory	Conversation Modeling
Healthcare	Medical Documentation	Reduced Paperwork	Specialized Terminology	Domain-Specific Models
Embedded Devices	Command Control	Hands-Free Operation	Battery Limitations	Model Optimization
Automotive	Driver Commands	Safety Enhancement	Acoustic Noise	Multi-Microphone Arrays
Accessibility	Assistive Technology	Independence	Diverse Speech Patterns	Specialized Training
Edge Computing	Hybrid Processing	Balanced Performance	Data Privacy	Split Computing

CONCLUSION

The history of voice AI technology can be seen as a spectacular intersection of progress in the fields of signal processing, machine learning, and computational linguistics, which has changed the basic nature of human-to-machine communication. This aspect will diminish the gap between human and synthetic speech as the technologies advance and generate more natural and intuitive paradigms of interaction, which opens accessibility to a wider range of user groups. They make up three core components, namely Voice Activity Detection, Transcription, and Text-to-Speech, which are different technical spaces where all three need to work in unison to achieve a harmonious voice experience, and any rise in any of the areas brings an increase in the effectiveness of the whole system. In the future, newer study areas such as multimodal integration, contextual comprehension, and emotional intelligence hold the promise of improving the potential of such systems even further. Voice AI exposes the need to work across disciplines that were traditionally disjoint, such as acoustics, linguistics, neural network design, and human-computer interaction. Effective processing methods and advanced neural structures enable Voice AI to shift more towards proactive interactions with conversation, and hence, a significant shift in the human-technological interactions will occur. The social implications of such developments go well beyond the convenience aspect to pose significant opportunities for inclusion coupled with productivity and newer ways of human expression via technology, making voice an important

Publisher: SARC Publisher

interface paradigm as one of the interfaces to the next generation computing system.

REFERENCES

1. Lee, S. "Advanced Signal Processing Techniques," Number Analytics, (2025). https://www.numberanalytics.com/blog/advanced-signal-processing-techniques-for-communications
2. Saxena, S. "10 Key Metrics You Need to Know for Measuring Voice AI Success," *Gnani.ai Resources*, (2025). https://www.gnani.ai/resources/blogs/10-key-metrics-you-need-to-know-for-measuring-voice-ai-success/
3. Beritelli, F., Casale, S., & Cavallaero, A. "A robust voice activity detector for wireless communications using soft computing." *IEEE Journal on Selected Areas in Communications* 16.9 (1998): 1818-1829.
4. Way With Words, "10 Metrics to Evaluate Performance of Speech Recognition Systems," Way With Words Resources. https://waywithwords.net/resource/speech-recognition-systems-performance/
5. Xiong, W., Droppo, J., Huang, X., Seide, F., Seltzer, M., Stolcke, A., ... & Zweig, G. "Achieving human parity in conversational speech recognition." *arXiv preprint arXiv:1610.05256* (2016).
6. Jurafsky, D. and Martin, J. H. "Speech and Language Processing," University of Colorado at Boulder, (2025). https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf
7. Ren, Y., Hu, C., Tan, X., Qin, T., Zhao, S., Zhao, Z., & Liu, T. Y. "Fastspeech 2: Fast and high-quality end-to-end text to speech." *arXiv preprint arXiv:2006.04558* (2020).
8. Zen, H., Senior, A., & Schuster, M. "Statistical parametric speech synthesis using deep neural networks." *2013 ieee international conference on acoustics, speech and signal processing*. IEEE, 2013.
9. Mathur, A., Lane, N. D., Bhattacharya, S., Boran, A., Forlivesi, C., & Kawsar, F. "Deepeye: Resource efficient local execution of multiple deep vision models using wearable commodity hardware." *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. 2017.
10. World Health Organization, & United Nations "Children's Fund. *Global report on assistive technology*. World Health Organization", (2022).

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:
Paul, V. "Voice AI Agents: The Technical Backbone of Modern Speech Interfaces" *Sarcouncil Journal of Multidisciplinary* 5.7 (2025): pp 610-617.

Publisher: SARC Publisher