

# Brain–Computer Interfaces Restoring Speech in Paralysis: Are We Near Clinical Reality?

Harshita Aggarwal<sup>1</sup>

Received January 17, 2026

Accepted April 30, 2026

Electronic access June 15, 2026

**Background/Objective:** Speech is a fundamental medium of human expression, yet neurological conditions such as amyotrophic lateral sclerosis (ALS), brainstem stroke, and locked-in syndrome can severely impair verbal communication while leaving cognition intact. Conventional assistive communication devices provide limited, effortful interaction, highlighting an urgent need for technologies that restore fluent speech. Brain–computer interfaces (BCIs) offer a novel solution by decoding neural activity directly from speech-related brain regions, bypassing damaged motor pathways. This review evaluates the current state of speech-restoring BCIs to determine their proximity to clinical reality and identify remaining challenges.

**Methods:** A structured narrative review was conducted following SANRA guidelines. Peer-reviewed studies from the past 10–15 years were identified through PubMed, IEEE Xplore, and Scopus. Inclusion criteria focused on human studies of BCIs aimed at speech restoration, utilizing intracortical or electrocorticography recordings, with measurable communication outcomes. Data extraction included participant characteristics, BCI modality, decoding strategy, performance metrics, and qualitative usability indicators. Narrative synthesis was employed due to heterogeneity in study design and outcomes.

**Results:** Invasive BCIs, particularly intracortical and high-density electrocorticography systems, have achieved low word error rates (9–24%), communication speeds approaching 60 words per minute, and near-real-time synthesized speech (latency < 100 ms) in small-scale invasive feasibility studies involving highly selected participants, typically with sample sizes of one to a few individuals, conducted under controlled experimental conditions with extensive system calibration and technical support. Non-invasive BCIs exhibit slower speeds and higher error rates but remain valuable alternatives. Emerging research on imagined speech decoding shows promises for effort-free communication. Longitudinal studies report stable performance and positive psychosocial impact, though small sample sizes, surgical invasiveness, and accessibility limit broader clinical adoption.

**Conclusions:** Speech-restoring BCIs have demonstrated functional, expressive communication under controlled conditions, marking a technological milestone. Widespread clinical deployment will require advances in safety, accessibility, and large-scale validation, alongside ethical safeguards for cognitive privacy.

**Keywords:** Brain–computer interface, speech restoration, neural decoding, ALS, locked-in syndrome, electrocorticography, imagined speech

## Introduction

### Background and Context

Speech is the primary medium through which humans convey thoughts, emotions, intentions, and social identity, making it essential for personal autonomy and social participation. Neurological conditions such as amyotrophic lateral sclerosis (ALS), brainstem stroke, traumatic brain injury, and advanced cerebral palsy can disrupt the motor pathways responsible for speech while leaving cognitive function intact, resulting in a profound dissociation between thought and expression. Individuals with locked-in syndrome, for example, may retain full awareness but be limited to minimal eye or facial move-

ments for communication, highlighting a major clinical and ethical challenge. Conventional augmentative and alternative communication (AAC) systems, including eye-tracking keyboards and switch-based interfaces, have improved access to communication but remain slow, cognitively demanding, and restrictive, limiting users to fragmented or effortful interactions rather than natural conversation<sup>1–4</sup>.

### Problem Statement and Rationale

Brain–computer interfaces (BCIs) provide a fundamentally different approach by bypassing damaged neuromuscular pathways and translating neural activity associated with speech planning or attempted articulation into text or synthesized voice. Although recent advances in invasive recording

<sup>1</sup> Pathways School, Noida, India

---

technologies such as intracortical electrode arrays and high-density electrocorticography (ECoG), combined with machine learning–based decoding, have enabled near-real-time speech synthesis with low latency (the time delay between neural signal input and generated output), current BCIs remain highly individualized, surgically invasive, and tested in small cohorts. Challenges including long-term signal stability, calibration, generalizability, emotional expressivity, and ethical concerns limit their broader clinical applicability<sup>1–4</sup>. In this review, we define clinical reality as the stage at which a speech brain–computer interface (BCI) can be safely, reliably, and practically used as a routine assistive communication technology for individuals who have lost the ability to speak due to neurological disease or injury. At this stage, the technology must move beyond controlled experimental demonstrations and function as a stable, therapeutic system capable of supporting everyday communication. A speech BCI is designed to detect neural signals from brain regions involved in speech planning and articulation, decode the intended linguistic content, and convert these signals into intelligible outputs such as synthesized speech or text in real time. The goal of this interface is to bypass damaged neuromuscular pathways that normally control the vocal tract, allowing individuals with conditions such as amyotrophic lateral sclerosis (ALS), brainstem stroke, or locked-in syndrome to communicate directly through neural activity, despite their inability to produce voluntary speech movements. These systems are clinically important because current assistive communication technologies such as eye-tracking keyboards and switch-based spelling interfaces typically allow only 5–20 words per minute, require sustained visual attention and motor control, and restrict users to slow, fragmented interactions that bear little resemblance to natural conversation. For a speech BCI to meet the threshold of clinical reality, several measurable criteria must be satisfied. First, the system must support communication speeds approaching those of normal conversation, generally estimated at around 50–70 words per minute, enabling spontaneous dialogue rather than slow, character-by-character message construction. Second, the system should achieve reliable decoding accuracy, usually defined as a word error rate (WER) the percentage of words decoded incorrectly below approximately 20–25% for open-vocabulary language decoding, ensuring that generated speech is intelligible without frequent manual correction. Third, the interface must demonstrate long-term stability and usability, with consistent performance maintained over at least 6–12 months of repeated use. This indicates that neural signals remain decodable and that recalibration demands can be managed outside tightly controlled research settings. Fourth, the technology must show reproducibility across multiple participants and independent research centers, demonstrating that high performance is not limited to a single optimized experimental case. Finally, a sys-

tem approaching clinical reality must present a clear pathway toward healthcare implementation, including acceptable surgical safety profiles for implanted devices, manageable maintenance requirements, and compatibility with regulatory processes governing clinical neurotechnology. Taken together, these benchmarks provide an operational framework for evaluating whether contemporary speech BCIs remain experimental prototypes or are truly approaching practical therapeutic deployment for individuals with paralysis<sup>3,5</sup>.

### **Significance and Purpose**

This review aims to critically evaluate the current state of speech-restoring BCIs to determine whether recent advances represent incremental improvements or a true transition toward clinical translation. By synthesizing performance, usability, safety, and long-term viability data, this study seeks to provide insight into how close these systems are to practical therapeutic use.<sup>1, 2,3</sup>

### **Objectives**

The objectives of this review are to (1) summarize recent technological and methodological advances in speech BCIs, (2) evaluate decoding performance and clinical feasibility, and (3) identify barriers and future directions for achieving routine clinical implementation.<sup>1–3</sup>

### **Scope and Limitations**

The review focuses exclusively on human studies of BCIs designed for speech restoration, including invasive and non-invasive systems. Non-speech BCIs and purely animal studies are excluded. Due to heterogeneity in study design, quantitative meta-analysis was not performed; instead, a narrative synthesis highlights trends, performance benchmarks, and translational potential.<sup>6–8</sup>

### **Theoretical Framework**

The study is grounded in systems of neuroscience, which posits that intended speech is encoded in distributed neural populations within motor and premotor cortices, providing the basis for neural decoding approaches<sup>9–13</sup>.

### **Methodology Overview**

A structured narrative review was conducted following SANRA guidelines, systematically identifying and evaluating peer-reviewed studies over the past 10–15 years. Key variables extracted included participant characteristics, BCI modality, decoding strategies, communication performance, and qualitative usability outcomes<sup>14–16</sup>.

---

## Related Work

### Advances in Neural Speech Neuroprostheses

In recent years, brain–computer interfaces (BCIs) have made remarkable progress in restoring speech for individuals who have lost the ability to communicate. Early pioneering studies by Anumanchipalli et al.<sup>4</sup> and Moses et al.<sup>3</sup> showed that neural activity recorded from the brain could be decoded into intelligible speech, providing the first clear demonstration that the brain’s signals could be translated directly into words. These studies were groundbreaking because they moved beyond the idea of simple neural control like moving a cursor or selecting letters to actually reconstructing naturalistic speech patterns.

Following this, more sophisticated systems emerged. Willett et al.<sup>17</sup> and Metzger et al.<sup>17</sup> developed high-performance BCIs capable of decoding continuous speech with relatively low error rates and impressive speed. These advancements suggest that practical, real-time communication might be possible for users with severe speech impairments. Meanwhile, Card et al.<sup>16</sup> addressed the critical issue of calibration, showing that neural decoding systems can be trained more efficiently, reducing the time and effort required for each user. Collectively, these studies highlight how far the field has come from proof-of-concept demonstrations to systems approaching functional usability.

Other researchers explored alternative methods for decoding neural signals. Angrick et al.<sup>18</sup> and Akbari et al.<sup>19</sup> examined approaches such as spectrogram reconstruction and mapping articulatory movements, which enhance both intelligibility and robustness. Similarly, Makin et al.<sup>6</sup> introduced encoder–decoder models that convert neural activity directly into text, improving efficiency and scalability. However, most of these systems still rely on invasive electrodes implanted in the brain and are tested primarily in laboratory settings. This raises questions about their readiness for broader clinical use, as real-world environments are far noisier and less predictable than controlled labs.

### Understanding How the Brain Encodes Speech

To build effective speech BCIs, it is crucial to understand how the brain represents speech. Mesgarani et al.<sup>20</sup> provided one of the most important insights, demonstrating that phonetic features are selectively encoded in the superior temporal gyrus. This discovery underpins phoneme-based decoding approaches, as it shows that the brain organizes speech information in a structured way that can be accessed for BCI applications. Complementing this, Tankus et al.<sup>11</sup> and Stavisky et al.<sup>9</sup> highlighted that speech-related information is distributed across both auditory and motor cortices, suggesting that speech processing involves multiple, interacting brain

regions.

Further research emphasized the role of the sensorimotor cortex in translating thought into speech. Khalighinejad et al.<sup>21</sup> and Wilson et al.<sup>10</sup> showed that neural activity in these regions corresponds closely to articulatory movements, supporting models that map brain signals directly to speech production mechanisms. Herff et al.<sup>7</sup> demonstrated that even complete phrases could be decoded from phoneme-level representations, highlighting the practical potential of this approach. Yet, despite these advances, there is no consensus on which brain regions or neural features are most reliable for decoding. Some studies suggest that motor signals are dominant, while others emphasize auditory or linguistic features. This uncertainty complicates model development and makes it difficult to directly compare different BCIs.

### Machine Learning and Generalization Challenges

Machine learning has been a key driver of progress in BCIs. Early work by Craik et al.<sup>22</sup> and Roy et al.<sup>23</sup> showed that deep neural networks, including convolutional and recurrent architectures, can uncover complex patterns in neural data, boosting classification accuracy. More recent models, including transformer-based and sequence-to-sequence architectures, have enabled more accurate decoding of continuous speech signals<sup>17</sup>. Despite these technological gains, generalization across users remains a major challenge. Most models are trained on data from a single participant and require time-intensive calibration, which limits scalability. Partial success has been reported in adapting models to new users, but Yuan et al.<sup>24</sup> found that performance often declines sharply when models encounter participants they were not trained on. This limitation is even more pronounced in non-invasive EEG-based imagined speech BCIs, where neural signals are noisier and more variable. Recent studies such as Rahman et al.<sup>25</sup>, Lopez-Bernal et al.<sup>26</sup>, Dhole et al.<sup>27</sup>, and Haresh & Begum<sup>28</sup> systematically reviewed methods for decoding imagined speech from EEG signals, highlighting the promise of deep learning in improving classification. Experimentally, Bhadra et al.<sup>29</sup> and Nguyen et al.<sup>30</sup> demonstrated systems where users could control BCIs through imagined speech, though their accuracy and reliability still lag behind invasive approaches.

These findings underscore a key tension: while machine learning offers powerful tools, achieving models that generalize across users remains a critical bottleneck.

### Non-Invasive and Imagined Speech Approaches

Non-invasive BCIs, particularly those based on EEG, offer a safer and more accessible alternative to invasive implants. Cooney et al.<sup>31</sup> and Nguyen et al.<sup>30</sup> showed that imagined

---

speech can be decoded from EEG signals, enabling basic communication without surgery. Panachakel & Ramakrishnan<sup>8</sup> reviewed these approaches extensively, highlighting how advances in signal processing and machine learning are gradually improving performance.

However, non-invasive systems face significant challenges. EEG has inherently low spatial resolution and is prone to noise, which limits decoding accuracy, especially for continuous or complex speech. Performance also varies across sessions and users, making consistent communication difficult<sup>25–27</sup>. While these systems are safer and more scalable, they currently cannot match the speed and accuracy of invasive BCIs, illustrating the trade-off between safety and performance.

### **Clinical Translation and Long-Term Stability**

Even high-performing BCIs face challenges when moving from the lab to the clinic. Signals recorded from implantable devices can degrade over time due to electrode wear, neural plasticity, or disease progression, requiring recalibration and limiting long-term usability<sup>32,33</sup>. Real-world applications also demand systems that are reliable, easy to use, and comfortable for patients<sup>5,34</sup>. Unfortunately, most studies remain short-term and lab-based, leaving questions about long-term stability largely unanswered.

### **Ethical and Societal Considerations**

As BCIs approach clinical and commercial deployment, ethical concerns become increasingly important. The ability to decode neural activity raises issues of mental privacy, informed consent, and data ownership. Ienca & Andorno<sup>35</sup> proposed the concept of “neurorights,” emphasizing legal protection for neural data, while Yuste et al.<sup>36</sup> argued for ethical frameworks to guide responsible neurotechnology development. These considerations are critical because errors, misuse, or unauthorized access to neural data could have profound consequences for users.

### **Synthesis of Gaps and Research Motivation**

Overall, speech BCIs have made extraordinary progress, yet several challenges remain. Invasive systems offer high accuracy but are limited in accessibility, whereas non-invasive EEG-based BCIs are safer but less precise. Generalization across users and long-term stability remain key obstacles, and methodological inconsistencies make comparing results difficult. Ethical concerns further complicate clinical deployment.

Taken together, these factors suggest that, while proof-of-concept demonstrations are impressive, the field still has a way to go before BCIs for speech restoration can be considered truly ready for widespread clinical use. This study aims to

examine whether recent advancements reflect genuine clinical readiness or primarily success under controlled experimental conditions.

## **Methods**

### **Review Guidelines**

This review followed the SANRA guidelines (Scale for the Assessment of Narrative Review Articles, a validated framework that assesses the quality of narrative reviews based on criteria such as clarity of objectives, comprehensiveness of the literature search, appropriate referencing, logical structure, and critical analysis) for narrative reviews to ensure systematic literature identification, transparent reporting, and methodological rigor. This approach ensured a systematic identification and evaluation of relevant literature despite the narrative review format.

### **Search Strategy**

A structured literature search was conducted across PubMed, IEEE Xplore, and Scopus to identify peer-reviewed studies investigating brain–computer interfaces (BCIs) for speech restoration. The search covered publications from 1 January 2010 to 15 October 2024, a period selected to capture the emergence and maturation of high-density neural recording technologies and machine learning–based speech decoding approaches<sup>37</sup>. The final search was performed on 15 October 2024, and no automated alerts or updates were applied thereafter. Searches were restricted to English-language publications.

Eligibility was limited to peer-reviewed journal articles and full-length conference papers with sufficient methodological detail to enable evaluation of study design, decoding approaches, and communication outcomes. Preprints, conference abstracts, editorials, and commentaries were excluded to maintain methodological rigor and ensure that only primary empirical evidence was synthesized. Review articles were not included in the primary analysis but were used selectively to provide background context and to identify relevant primary studies. Reference lists of included studies were additionally screened manually to identify relevant articles not captured in the database search. Search strategies combined controlled vocabulary terms (e.g., MeSH terms in PubMed where applicable) and free-text keywords related to brain–computer interfaces, neural speech decoding, and communication impairment, using Boolean operators (AND/OR) and truncation where supported. Field tags were applied where appropriate (e.g., title/abstract fields) to improve search specificity.

---

An example search string used in PubMed was

("brain-computer interface"[Title/Abstract] OR "BCI"[Title/Abstract] OR "neural decoding"[Title/Abstract] OR "speech neuroprosthesis"[Title/Abstract]) AND ("speech"[Title/Abstract] OR "speech decoding"[Title/Abstract] OR "speech synthesis"[Title/Abstract] OR "communication"[Title/Abstract]) AND ("paralysis"[Title/Abstract] OR "locked-in syndrome"[Title/Abstract] OR "amyotrophic lateral sclerosis"[Title/Abstract] OR "ALS"[Title/Abstract] OR "anarthria"[Title/Abstract]) Equivalent search structures were adapted for IEEE Xplore and Scopus using database-specific syntax, indexing systems, and field restrictions to ensure consistency in search scope across platforms.

### Study Management and Screening

All retrieved records were managed using EndNote X9 (Clarivate) for de-duplication and organization. Titles and abstracts were screened for relevance to speech-related BCIs, followed by full-text evaluation of articles that met the initial criteria. To ensure accuracy and minimize bias, data extraction and eligibility assessment were carefully verified against original reports, and key details were cross-checked within the full texts. Reference lists of included articles were also manually screened to identify additional relevant studies not captured in the database search. Screening, eligibility assessment, and data extraction were conducted by a single reviewer (the author). To enhance accuracy and reduce the risk of extraction errors, all included studies underwent a secondary verification step in which extracted data and eligibility decisions were rechecked against the full-text articles.

This process involved repeated cross-referencing of study characteristics, methodological details, and reported outcomes with the original publications. No independent second reviewer was involved in the screening or extraction process. As a result, there is a potential risk of selection bias or missed studies, which is an inherent limitation of single-reviewer narrative reviews. However, this risk was mitigated through the use of predefined inclusion and exclusion criteria, consistent application of search strategies across databases, and systematic full-text verification of all included studies. Eligibility and data assessments were carefully cross-checked against full texts to ensure accuracy and minimize bias.

### Inclusion and Exclusion Criteria

#### *Studies were included if they*

Studies were included if they were peer-reviewed original research articles, involved human participants with paralysis or severe speech impairment, investigated brain-computer interfaces explicitly aimed at speech production, decoding, or

synthesis (including both text and voice output), and reported empirical outcomes related to communication performance.

#### *Studies were excluded if they*

Studies were excluded if they were conducted exclusively in animal models, focused on non-speech brain computer interfaces such as cursor control or limb movement, were reviews, editorials, or conference abstracts lacking sufficient methodological detail, or failed to provide an adequate description of decoding methods or outcome measures.

Missing or incomplete data were noted, and study authors were contacted when feasible to clarify key methodological or outcome details.

For the purposes of this review, included studies were categorized into two analytically distinct groups to ensure clarity in interpreting translational relevance. The first category comprised target clinical population studies, defined as investigations involving individuals with severe speech impairment or loss of voluntary speech due to neurological conditions such as amyotrophic lateral sclerosis (ALS), locked-in syndrome, brainstem stroke, or anarthria. These studies provide direct empirical evidence regarding the capacity of brain-computer interfaces to restore functional communication in the intended clinical population. The second category comprised foundational speech decoding studies, defined as investigations conducted in participants without primary speech impairment, most commonly individuals undergoing intracranial monitoring for clinical purposes (e.g., epilepsy).

In these studies, neural recordings are used to characterize the encoding of speech production and perception. While such studies provide critical mechanistic and algorithmic insights, such as identifying neural correlates of articulatory planning, optimizing decoding architectures, and validating signal acquisition methods, they do not constitute direct evidence of clinical speech restoration. This distinction was applied consistently during data extraction and narrative synthesis. Findings from foundational studies were used to inform mechanistic understanding and technological feasibility, whereas conclusions regarding clinical performance, usability, and translational applicability were derived primarily from studies involving the target clinical population. Although study selection followed a predefined, multi-stage screening process including database identification, duplicate removal using EndNote X9 (Clarivate), title and abstract screening, and full-text eligibility assessment formal numerical tracking of records at each stage was not systematically documented, as the review was designed and conducted in accordance with SANRA guidelines for narrative reviews rather than PRISMA-based systematic review protocols.

To maintain methodological rigor despite the absence of a formal flow diagram, all retrieved records were screened

---

against explicit, predefined inclusion and exclusion criteria, and eligibility decisions were verified through full-text evaluation to ensure consistency and relevance. Reference list screening was additionally performed to minimize the risk of missing pertinent studies. While this approach supports transparency in study selection, the absence of stage-wise numerical reporting limits full reproducibility of the screening flow. Future work may incorporate prospective PRISMA-style tracking to provide a more detailed quantitative account of study identification, screening, and inclusion.

### **Data Extraction**

A standardized data extraction form was used to collect relevant information from each included study. Extracted variables included study characteristics, such as year of publication, study design, and sample size; participant information, including neurological condition, degree of paralysis, and severity of speech impairment; and BCI characteristics, such as recording modality (e.g., intracortical arrays or electrocorticography) and targeted brain regions. Additionally, details on decoding and output methods were recorded, including text-based decoding, articulatory modeling, and speech synthesis approaches. Finally, outcome measures were extracted, encompassing quantitative metrics such as communication speed, accuracy, word error rate, and latency, as well as qualitative assessments of usability and speech intelligibility.

### **Data Synthesis**

Due to heterogeneity in study designs, participant populations, and outcome metrics, a narrative synthesis was performed. Findings were organized thematically and grouped according to BCI technology, decoding strategy, and stage of clinical translation (proof-of-concept, feasibility studies, or early clinical trials). This approach allowed for comparison of methodological trends, performance benchmarks, and translational progress while acknowledging differences in experimental context. Due to the heterogeneity of BCI technologies, participant populations, and outcome metrics, quantitative meta-analysis was not feasible; instead, a narrative synthesis was performed to identify thematic trends and translational progress. To maintain interpretive clarity, results were synthesized and reported separately for target clinical population studies and foundational speech decoding studies.

### **Quality Assessment and Risk of Bias**

The methodological quality of included studies and potential sources of bias were assessed using a structured qualitative framework specifically developed for this review, ensuring consistency and transparency in evaluation. Each study was systematically examined across several predefined domains

relevant to translational brain–computer interface (BCI) research. These domains included sample size and participant characteristics, with particular attention to whether studies involved target clinical populations or non-clinical participants; study duration and follow-up, distinguishing between single-session experiments and longitudinal assessments of performance stability; and evaluation context, specifically whether decoding performance was assessed in real-time (online) conditions or through offline analyses. In addition, the completeness and clarity of outcome reporting were evaluated. This included examining whether studies consistently reported key performance metrics such as word error rate (WER), communication speed, latency, and measures of intelligibility. Consideration was also given to reproducibility and generalizability, including the number of participants, the extent of cross-session validation, and whether decoding models were tested on novel or unconstrained linguistic inputs rather than fixed training sets.

Finally, technical robustness and device stability were assessed based on reported signal consistency over time, calibration requirements, and evidence of sustained usability across sessions. Rather than applying a numerical scoring system, each domain was assessed descriptively to account for the substantial heterogeneity in study design, methodologies, and reporting standards. This approach allowed for a nuanced interpretation of study strengths while avoiding inappropriate quantitative comparisons across fundamentally different experimental paradigms.

Overall, the body of evidence was characterized by small sample sizes, limited long-term follow-up, and variability in reporting practices all factors that were carefully considered when interpreting findings related to clinical feasibility and translational readiness. Qualitative outcomes, including patient-reported psychosocial impact and safety observations, were synthesized using a structured interpretive narrative approach, reflecting the substantial heterogeneity in reporting formats, outcome definitions, and follow-up measures across studies. Due to the absence of standardized and consistently reported metrics for patient-reported outcomes and adverse events, a formal quantitative extraction or comparative framework was not feasible. Consequently, these findings are presented as qualitative trends, derived from study-level observations rather than systematically aggregated data.

### **Scope of the Review**

The scope of this review was clearly defined to focus exclusively on human studies of speech-related BCIs, evaluating translational potential, performance benchmarks, and technological innovations relevant to clinical implementation, while excluding non-speech BCI studies to maintain focus on communication restoration.

---

## Results

### Overview of Included Studies

The body of research examining brain–computer interfaces (BCIs) for speech restoration is relatively recent, reflecting both the technical difficulty of decoding speech-related neural signals and the ethical and clinical constraints associated with working with individuals who have severe paralysis. The majority of studies included in this review rely on invasive neural recording techniques, particularly intracortical microelectrode arrays and high-density electrocorticography (ECoG). These approaches directly measure neural activity from regions of the brain involved in speech production, most commonly the speech motor cortex and nearby premotor regions, which are responsible for planning and executing articulatory movements<sup>1–4</sup>. Participants across these studies were predominantly adults diagnosed with conditions such as amyotrophic lateral sclerosis (ALS), brainstem stroke, or locked-in syndrome. These individuals typically retain full cognitive function but experience profound speech impairment or complete loss of voluntary speech<sup>1,3</sup>. Because these studies involve invasive procedures and extensive calibration, sample sizes are necessarily small, usually ranging from one to five participants per study<sup>1–3</sup>.

The studies reviewed span approximately 2010 to 2025, a period marked by rapid advancements in neural recording hardware, machine-learning-based decoding algorithms, and real-time speech synthesis technologies. Early investigations focused primarily on enabling basic communication, such as selecting letters sequentially to spell words or making binary choices<sup>7,14</sup>. In contrast, more recent studies demonstrate the ability to decode continuous speech and generate real-time synthesized voice output<sup>1,2,4,6</sup>.

Study durations ranged from single laboratory sessions to multi-week or multi-month longitudinal assessments<sup>1,2</sup>. Importantly, several longitudinal studies report stable performance over extended periods, suggesting that these systems may be feasible for repeated and sustained real-world use<sup>1,2</sup>. Overall, the literature reflects significant technological progress and increasing clinical relevance, while also highlighting persistent limitations related to participant diversity and small cohort sizes. These studies comprise both investigations conducted in target clinical populations with severe speech impairment and foundational studies involving participants without primary speech deficits, which are distinguished in the following sections to clarify their respective contributions to translational progress. To avoid conflating distinct paradigms, results are organized into (1) attempted speech decoding in clinical populations, (2) foundational overt speech decoding studies in non-paralyzed participants, and (3) imagined (inner) speech as a proof-of-concept paradigm.

### Attempted Speech Decoding in Clinical Populations (Paralysis / Locked-in Syndrome)

This section focuses on studies involving individuals with severe speech impairment, which provide the primary basis for evaluating clinical feasibility and translational readiness. In studies of attempted speech decoding in individuals with paralysis, a central indicator of success is decoding performance, typically evaluated using metrics such as word error rate, communication speed, intelligibility, and latency. In several landmark intracortical studies, participants with amyotrophic lateral sclerosis (ALS) achieved low word error rates under clearly defined experimental conditions. In the study by Willett et al. (2023), a single participant ( $n = 1$ ) with ALS performed attempted speech tasks, achieving a word error rate of approximately 9% from a constrained vocabulary of , 50 words under a closed-set, supervised classification paradigm. In the same participant, decoding from a large open vocabulary (, 125, 000 words) yielded a higher word error rate of approximately 24% under an open-vocabulary language model–integrated neural decoding framework evaluated in an online setting. This system combined neural signal decoding with probabilistic language modeling to improve sequence prediction. These results were obtained following extensive participant-specific training over multiple sessions, and therefore reflect optimized, single-subject performance rather than generalized clinical outcomes.

In addition to accuracy, communication speed is a critical determinant of clinical utility. These estimates must be interpreted within tightly defined experimental contexts. In the intracortical study by Willett et al. (2023), a single participant ( $n = 1$ ) with ALS achieved approximately 60 words per minute during online decoding of attempted speech, using an open-vocabulary neural decoder integrated with a probabilistic language model.

Similarly, Metzger et al. (2023) reported comparable communication speeds (, 60 words per minute) in a single participant ( $n = 1$ ) with ALS using ECoG-based recordings during attempted speech-based text generation tasks involving structured phrase production. In both cases, performance was measured under highly controlled experimental conditions, including repeated calibration sessions, constrained conversational prompts, and participant-specific model optimization. As such, these communication rates reflect upper-bound performance under optimized laboratory conditions, rather than unconstrained, real-world conversational use, although these values approach the lower range of natural conversational speech.<sup>1,2</sup> Latency measurements are similarly dependent on decoding architecture and evaluation conditions. Notably, these systems were capable of generalizing to novel phrases, indicating that decoding algorithms are learning underlying speech representations rather than memorizing fixed

---

outputs<sup>1,6</sup>.

By contrast, non-invasive BCIs, particularly EEG-based systems, typically achieve communication speeds of approximately 15–20 words per minute in spelling-based or selection-based paradigms, with substantially higher and more variable error rates due to low spatial resolution, signal attenuation through the skull, and susceptibility to noise artifacts<sup>8</sup>. These systems are generally evaluated in controlled laboratory settings using predefined tasks, and while safer, they currently lack the signal fidelity required for high-speed, continuous speech decoding.

Decoding strategy also influences performance. Text-based decoding, such as letter-by-letter spelling, tends to be more reliable but inherently slower due to its sequential nature<sup>7,10</sup>. In contrast, direct neural-to-speech synthesis approaches have demonstrated the potential for faster and more natural communication, but their performance in individuals with severe paralysis remains less extensively validated and requires further investigation. Advances in deep learning have substantially improved intelligibility, reduced word error rates, and enabled the production of syntactically coherent speech across decoding approaches<sup>1,6</sup>.

### **Foundational Speech Decoding Studies in Non-Paralyzed Populations (Overt Speech Tasks)**

In a foundational study conducted in participants undergoing clinical epilepsy monitoring (non-paralyzed population performing overt speech tasks), Anumanchipalli et al. (2019) demonstrated neural-to-speech synthesis from overt speech tasks with latencies below 100 milliseconds using ECoG recordings from participants undergoing clinical epilepsy monitoring. The system employed a two-stage decoding pipeline, mapping neural activity to articulatory representations and subsequently to acoustic speech via a vocoder. These latency values were obtained in controlled laboratory settings with offline-trained models deployed in near real-time, and therefore do not represent communication performance in clinical populations with paralysis. Therefore, while sub-100 ms latency demonstrates the technical feasibility of real-time speech synthesis, its clinical generalizability remains constrained by differences in participant population and experimental design, rather than fully autonomous clinical communication, despite enabling near-continuous verbal interaction in controlled settings<sup>1,4</sup>.

Speech-restoring the brain–computer interfaces (BCIs) encompass several distinct decoding paradigms that differ fundamentally in their computational frameworks, neural signal requirements, linguistic flexibility, and implications for clinical translation. These paradigms include small-vocabulary classification systems, open-vocabulary text decoding approaches, and neural-to-speech synthesis models, each repre-

senting a different stage in the progression from constrained communication toward naturalistic speech restoration. Small-vocabulary classification systems are designed to map neural activity onto a fixed, predefined set of linguistic outputs, typically consisting of a limited number of words, phonemes, or commands. These systems often rely on supervised classification algorithms trained to recognize repeated neural patterns associated with specific targets. Because the output space is restricted, classification boundaries are well-defined, allowing these systems to achieve relatively low error rates under controlled experimental conditions. However, their clinical utility is inherently limited: users are unable to generate novel or spontaneous utterances beyond the predefined vocabulary, restricting communication to selection-based interactions rather than fully generative language use. By contrast, open-vocabulary text decoding approaches aim to translate neural activity into unconstrained linguistic sequences, enabling users to construct arbitrary words and sentences. These systems typically employ sequence-to-sequence models or encoder–decoder architectures that map continuous neural signals onto probabilistic language representations. Unlike classification-based systems, open-vocabulary decoding must resolve ambiguity across a vastly larger linguistic space, requiring the integration of neural features with statistical language models to preserve coherence and grammatical structure. While this approach provides substantially greater linguistic flexibility, it also introduces increased computational complexity and typically results in higher word error rates compared with constrained systems, particularly in spontaneous or unconstrained communication contexts.

The ability to generate novel sentences represents a critical requirement for clinically meaningful communication. Neural-to-speech synthesis models build on previous approaches by attempting to reconstruct continuous acoustic speech signals directly from neural activity, rather than producing intermediate text representations. These systems typically model articulatory or acoustic features of speech and then synthesize audible output using vocoder-based or neural speech generation frameworks. Because speech production involves rapid, temporally precise coordination of multiple articulators, these models require high-resolution neural recordings, such as intracortical signals or high-density electrocorticography, to capture the fine-grained dynamics of motor speech. This paradigm offers the potential for more naturalistic and expressive communication, including prosody, intonation, and speech timing. However, it also introduces additional challenges, including maintaining signal fidelity, managing model training complexity, and ensuring robustness across recording sessions. Importantly, these paradigms are not directly comparable on a single performance scale. Differences in vocabulary size, output modality, and decoding strategy impose fundamentally different constraints on accuracy,

---

speed, and usability. For instance, low error rates reported in small-vocabulary systems cannot be interpreted as equivalent to performance in open-vocabulary decoding, where the combinatorial space of possible outputs is substantially larger. Similarly, communication rates achieved in text-based systems may not directly translate to neural-to-speech synthesis, where latency, intelligibility, and acoustic quality introduce additional dimensions of evaluation. Therefore, meaningful assessment of translational progress requires that each paradigm be evaluated within its respective methodological context, with explicit consideration of underlying assumptions, performance metrics, and clinical objectives, rather than being aggregated into a single generalized narrative of advancement.

Accordingly, all reported performance metrics including word error rate, communication speed, and latency are highly contingent on study-specific parameters, including participant population (e.g., ALS vs epilepsy), evaluation setting (online vs offline), decoding paradigm (closed-set classification vs open-vocabulary generation vs neural-to-speech synthesis), and the use of external language models. Without explicit alignment of these factors, cross-study comparisons risk overstating performance and obscuring meaningful differences in system capability and translational readiness.

For table 1A and 1B- Taken together, the study-level comparisons presented in Table 1A and Table 1B demonstrate that reported performance across speech-restoring BCIs is highly dependent on decoding paradigm, recording modality, and vocabulary scale. Differences in methodological design, particularly the use of constrained versus open vocabularies and text-based versus neural-to-speech decoding approaches, introduce fundamentally different performance constraints, limiting direct comparability across studies. These findings underscore the necessity of evaluating each system within its specific technical and experimental context when assessing translational progress toward clinical implementation.

Table 2 presents an aggregated comparison of speech BCI performance across major recording modalities, highlighting clear differences in decoding capability, communication speed, and latency. Invasive systems, particularly intracortical approaches, demonstrate the highest performance, achieving low word error rates and communication speeds approaching conversational levels, alongside near real-time output. These results reflect the high spatial and temporal resolution of direct cortical recordings, which enable detailed capture of speech-related neural activity. In contrast, non-invasive systems exhibit substantially lower communication speeds and higher variability in accuracy, largely due to reduced signal fidelity and increased noise in scalp-recorded data. However, these modality-level comparisons must be interpreted with caution. As demonstrated in Tables 1A and 1B, performance outcomes are strongly influenced not only by recording modality but

also by study-specific factors, including vocabulary size, decoding paradigm (e.g., fixed classification vs open-vocabulary decoding), and task design. For example, low word error rates observed in small-vocabulary systems are not directly comparable to those reported in large-scale open-vocabulary decoding, where linguistic complexity is significantly greater.

Furthermore, the high performance of invasive systems is derived from studies with extremely small sample sizes (typically  $n = 1-5$ ), conducted under controlled experimental conditions with extensive calibration and technical support. The requirement for neurosurgical implantation also limits scalability and clinical accessibility. Consequently, while Table 2 highlights the current performance ceiling of speech BCIs, it does not fully capture the translational constraints that determine real-world clinical viability.

### **Imagined (Inner) Speech Decoding as a Proof-of-Concept Paradigm**

Imagined (inner) speech decoding represents an early-stage, proof-of-concept paradigm that is scientifically and clinically distinct from both attempted speech in individuals with paralysis and overt speech decoding in non-paralyzed populations. Recent research has begun to explore the possibility of decoding inner speech the words a person imagines without attempting physical articulation. Several studies have shown that imagined speech evokes neural patterns similar to attempted speech, and these patterns can be decoded using machine learning models, although with lower signal strength and higher error rates compared to attempted speech recordings<sup>7,8</sup>. In one proof-of-concept demonstration, researchers reported up to approximately 74% decoding accuracy for imagined sentences. However, this performance was achieved under highly constrained experimental conditions, involving predefined tasks and controlled evaluation settings. Many of these studies also rely on offline or semi-online analysis pipelines rather than fully continuous, real-time communication. These findings highlight the potential of imagined speech decoding as a long-term avenue for more natural, effort-free communication, yet current systems remain at an early proof-of-concept stage and are not yet suitable for reliable clinical use in individuals with severe paralysis. While this emerging paradigm could eventually enable communication without requiring overt motor effort or articulatory attempts, substantial challenges remain. These include low signal reliability, limited generalizability across participants and contexts, and ethical considerations, such as ensuring that only intended thoughts are decoded. Accordingly, results from imagined speech studies should not be interpreted as indicative of near-term clinical translation for individuals with severe speech impairment.

**Table 1A** Study characteristics of key speech BCI investigations

Study (year)	Diagnosis	n	Modality	Brain region	Task type	Vocabulary
Willett et al. (2023)	ALS	1	Intracortical	Motor cortex	Open-vocabulary text decoding	125, 000
Metzger et al. (2023)	ALS	1	ECoG	Speech motor cortex	Text decoding (phrases)	1, 000
Moses et al. (2021)	Anarthria (stroke)	1	ECoG	Ventral motor cortex	Fixed phrase classification	50
Anumanchipalli et al. (2019)	Epilepsy	Multiple	ECoG	Speech cortex	Neural-to-speech synthesis	Sentences

**Table 1B** Performance and clinical outcomes of key studies

Study (year)	WER/Accuracy	Speed	Latency	Follow-up	Key constraints
Willett et al. (2023)	24% WER	60	<100 ms	Months	Single participant; invasive; extensive training
Metzger et al. (2023)	15–25% WER	60	<200 ms	Weeks	Small sample; short duration; surgical implantation
Moses et al. (2021)	20–30% WER	15–20	Weeks	Weeks	Limited vocabulary; slower communication
Anumanchipalli et al. (2019)	Intelligibility based	Real time	Short	Short	Not target population; partial offline validation

**Table 2** Performance comparison across BCI modality

BCI Type	Vocabulary Size	Word error rate	Communication speed	Latency
Invasive (intracortical)	, 50 words	, 9 (reported in controlled studies)	60-65	<100 ms (speech synthesis systems)
Invasive (intracortical)	, 125, 000 words	, 24 (Large vocabulary decoding)	, 60	<100 ms (speech synthesis systems)
Non Invasive (EEG)	Limited	Higher (variable)	, 15-20	Typically, higher

### Comparison of BCI Approaches

Speech-restoring BCIs can be broadly classified into invasive and non-invasive systems, each presenting distinct advantages and limitations relevant to clinical translation. Invasive BCIs, including intracortical microelectrode arrays and ECoG grids, provide high spatial and temporal resolution, allowing precise capture of neural signals associated with articulatory movements. These systems consistently achieve the highest decoding accuracies and fastest communication speeds, enabling near-real-time speech output<sup>1,2,4</sup>. However, they require neurosurgical implantation, which introduces risks such as infection and raises concerns regarding long-term signal stability<sup>1</sup>.

Non-invasive BCIs, such as EEG-based systems, eliminate surgical risk and are easier to deploy in clinical settings. However, they suffer from lower signal fidelity, which results in

slower communication and increased error rates<sup>8</sup>. As a result, these systems currently fall short of enabling natural conversational speech. Emerging hybrid and semi-invasive approaches aim to balance safety and performance by improving signal quality while minimizing invasiveness<sup>14,15</sup>. Decoding strategy further differentiates system performance. Direct speech synthesis enables rapid, fluent output but requires extensive training and high-quality neural data<sup>1,4</sup>. Text-based decoding, while slower and less expressive, remains more robust under limited signal conditions<sup>7</sup>. At present, invasive BCIs offer superior performance, while non-invasive approaches hold promise for broader applicability as decoding techniques improve.

---

## Clinical Translation Indicators

Indicators of progress toward clinical reality include long-term usability, safety, and patient acceptance. Longitudinal evidence shows that participants can use invasive speech BCIs for several hours per day across multiple weeks or months while maintaining stable decoding performance<sup>1,2</sup>. These findings suggest that such systems can support meaningful real-world communication, including composing messages, engaging in conversation, and expressing personal thoughts. Beyond quantitative performance, qualitative feedback from participants highlights the impact on daily life and personal autonomy. Users consistently report that the ability to generate speech, even with some errors or slower speed, significantly reduces frustration and reliance on caregivers, allowing for more spontaneous and natural social interaction<sup>1,3</sup>. Many participants describe a sense of regained agency and personal identity, noting that being able to express tone, emotion, or intent is psychologically meaningful. These observations illustrate that technical metrics such as word error rate and communication speed translate directly into practical, lived benefits, providing insight into how BCIs could function as real-world clinical tools.

Safety remains a central concern, particularly for invasive systems. Although surgical implantation is required, reported adverse events across studies are minimal, typically involving minor wound complications or temporary calibration difficulties<sup>1,3</sup>. Despite these promising safety profiles, the invasive nature of the systems inherently limits widespread accessibility, particularly for individuals with limited healthcare resources, comorbidities, or in regions without specialized neurosurgical support. Participant acceptance in early trials has been largely positive. Users report increased independence, reduced frustration associated with communication barriers, and improved social interaction<sup>1,3</sup>. Importantly, the benefits extend beyond functional communication to include emotional and psychological well-being. However, accessibility challenges remain a significant barrier to translating these systems into broad clinical reality. The complexity, cost, and need for trained clinicians to perform implantation, calibration, and long-term maintenance restrict availability, meaning that only a small subset of patients can realistically access these technologies at present.

Nonetheless, larger and more diverse studies are necessary to fully assess long-term comfort, adherence, and psychological impact across broader patient populations. Qualitative insights suggest that for clinical reality to be fully achieved, future systems must address not only technical performance and safety but also equitable access, usability in everyday environments, and long-term support infrastructure.

## Qualitative Outcomes and Real-World Communication Challenges

Beyond quantitative performance metrics, several studies reported qualitative findings that directly reflect the real-world communication challenges faced by individuals with severe paralysis. Participants consistently described profound limitations associated with existing assistive communication methods, such as eye-tracking or switch-based spelling systems, which are often slow, cognitively demanding, and restrict spontaneous conversation<sup>1,3</sup>. In this context, speech-restoring BCIs were reported to address a critical unmet need: the ability to communicate fluidly and expressively in everyday situations<sup>1,2</sup>. Users noted that even partial restoration of spoken output such as synthesized or reconstructed speech enabled more natural social interaction, reduced reliance on caregivers for mediated communication, and increased autonomy in daily decision-making<sup>3</sup>. Qualitative reports further emphasized that the value of these systems extended beyond message transmission, affecting emotional well-being by allowing participants to convey tone, intent, and personal identity. Although derived from small cohorts, these observations demonstrate that the technical capabilities measured in decoding accuracy and speed correspond to meaningful improvements in real-world communication, directly addressing the core problem of social and functional isolation experienced by people who have lost the ability to speak.

## Discussion

### Summary of Key Findings

This review synthesizes over a decade of research on speech-restoring brain-computer interfaces (BCIs), explicitly distinguishing between two types of studies: (1) investigations conducted in clinical populations with paralysis or severe speech impairment, and (2) foundational speech decoding studies conducted in non-target populations. Evidence from clinical studies shows that invasive BCIs, particularly intracortical and high-density electrocorticography (ECoG) systems, can achieve low word error rates, high communication speeds, and near-real-time speech synthesis under controlled experimental conditions, albeit in small participant cohorts<sup>1, 2,3</sup>. In parallel, foundational studies involving individuals without severe speech impairment have contributed critical advances in decoding methodology, neural feature extraction, and speech representation, enabling improvements in model generalization and continuous speech synthesis<sup>4,6,18,38</sup>. While these studies are essential for technological development, they do not directly demonstrate clinical communication restoration. Non-invasive BCIs, though safer and more accessible, currently demonstrate lower decoding accuracy and reduced re-

---

liability due to limited spatial resolution and greater susceptibility to noise, particularly in imagined speech paradigms<sup>31,39</sup>. These systems remain important alternatives for individuals who cannot undergo neurosurgical implantation, but they do not yet achieve performance comparable to invasive systems in clinical contexts. Research on inner (imagined) speech decoding, largely conducted in non-clinical populations, represents an important foundational direction<sup>40,41</sup>. However, this work remains at an early stage and should be interpreted as a long-term research trajectory, rather than as evidence of current clinical feasibility. Collectively, these findings indicate that speech BCIs have made substantial technical progress. Nevertheless, conclusions regarding clinical reality must be grounded primarily in evidence derived from target patient populations.

### **Implications for Clinical Practice**

#### **Task and Population Distinction in Speech BCI Evidence**

##### **Attempted and Overt Speech Decoding in Clinical Populations**

From a clinical standpoint, the most significant implication of the reviewed literature is that speech BCIs are approaching functional viability for carefully selected patient populations with paralysis or severe speech impairment, based on a small number of clinical studies conducted under controlled experimental conditions. Invasive systems have demonstrated sustained performance over periods of weeks to months in clinical participants, maintaining relatively stable decoding accuracy and acceptable short-term safety profiles<sup>1-3</sup>. Communication speeds approaching approximately 60 words per minute, combined with low latency and intelligible speech synthesis, suggest that these systems may support aspects of everyday conversational interaction, at least in structured settings<sup>1-3</sup>. Importantly, these findings are derived specifically from studies involving individuals with paralysis or severe speech impairment and should be clearly distinguished from foundational speech decoding studies conducted in non-target populations. Despite these advances, current readiness remains limited to highly controlled research or clinical environments, and to a small number of participants who have access to extensive technical and medical support. Substantial calibration, specialist oversight, and ongoing system maintenance are still necessary. As such, while speech BCIs may be considered clinically realistic for experimental or compassionate-use scenarios, they are not yet suitable for routine, large-scale integration into standard healthcare practice.

Equally important are the psychological and social benefits reported by participants in clinical studies involving individuals with severe speech impairment, including perceived increases in autonomy, reductions in communication-related frustration, and partial restoration of personal identity, as re-

ported in a small number of participants (typically  $n = 1-5$ ) under controlled, researcher-supported experimental conditions<sup>1-3</sup>. These findings are derived exclusively from small clinical cohorts and are not informed by foundational speech decoding studies conducted in non-target populations. These outcomes suggest that speech BCIs may address not only functional communication deficits but also broader aspects of mental well-being and quality of life, underscoring the need to incorporate patient-reported outcomes into clinical evaluation frameworks.

### **Foundational Speech Decoding Studies (Non-Clinical Populations)**

Foundational speech decoding research has played a critical role in the development of contemporary speech BCI systems. These studies primarily involve overt speech production in non-clinical participants, and are conducted in controlled experimental settings to understand how neural activity encodes speech-related information. Advances in neural recording methods, signal processing, and decoding algorithms have been informed by foundational neural decoding research, including studies such as Angrick et al. and Chartier et al. These studies have demonstrated that speech-related cortical activity can be decoded into phonemes, words, or articulatory features under controlled conditions, forming the basis for more advanced speech reconstruction systems. In addition, similar experimental constraints, including small sample sizes, have been observed in foundational intracortical decoding studies such as Pandarinath et al. reflecting the technical and methodological complexity of high-resolution neural recording. Importantly, while these studies provide essential methodological and theoretical foundations, they do not involve individuals with paralysis or severe speech impairment and therefore do not constitute direct evidence of clinical translation. Instead, they should be understood as enabling technologies that inform the development of clinically oriented speech BCI systems.

### **Imagined/Inner Speech Decoding (Proof-of-Concept)**

Imagined or inner speech decoding represents a distinct and emerging line of research within the broader field of speech BCIs. Unlike overt or attempted speech paradigms, these approaches aim to decode internally generated speech representations in the absence of actual articulation or motor output. Current work in this domain, including studies such as Panachakel & Ramakrishnan, remains at a proof-of-concept stage and is primarily conducted in non-clinical or highly controlled experimental contexts. These studies demonstrate the theoretical possibility of decoding internally generated linguistic content from neural activity; however, performance re-

---

mains substantially lower and less reliable than that observed in overt or attempted speech paradigms. A key challenge in imagined speech decoding is distinguishing intentional communication from spontaneous or background neural activity, particularly in unconstrained or real-world settings. This limitation introduces significant technical and ethical complexities, including concerns related to decoding specificity, user control, and cognitive privacy. At present, imagined speech decoding has not been validated in clinical populations with severe speech impairment and does not yet support reliable, real-time communication. As such, it should be regarded as an early-stage research direction with long-term potential, rather than a technology approaching clinical implementation.

### *Rehabilitation and Psychosocial Impact*

Beyond communication restoration, speech BCIs may also hold rehabilitative potential in clinical populations with paralysis or partial speech impairment. Repeated engagement of speech motor networks may contribute to neural plasticity (the brain's ability to reorganize and form new connections), particularly in individuals with partial speech impairment. This raises the possibility that BCIs could complement conventional speech therapy rather than solely compensating for lost function. However, this rehabilitative potential remains largely theoretical and is supported primarily by limited clinical observations rather than systematic longitudinal evidence.

### *Surgical Risk and Autonomy*

A central ethical concern in high-performance speech BCIs, particularly in clinical applications involving individuals with paralysis or severe speech impairment, is the requirement for neurosurgical implantation. Although adverse events reported to date in small clinical cohorts have been limited, invasive procedures inherently carry risks such as infection, inflammation, and long-term device stability issues<sup>1,2,3</sup>. For individuals with progressive neurodegenerative conditions, such as amyotrophic lateral sclerosis (ALS), these risks raise difficult ethical questions regarding the balance between surgical burden and anticipated long-term benefit. Additionally, the semi-permanent nature of implanted devices introduces concerns around informed consent, long-term autonomy, and the feasibility of device removal or replacement if technologies fail or become obsolete. Current speech BCI systems rely on complex hardware–software integration, including implanted electrodes, neural signal acquisition pipelines, and real-time decoding algorithms. Each of these components represents a potential point of failure for example, signal degradation, electrode malfunction, software instability, or decoding breakdown. In the event of device malfunction, users may experience loss of decoding accuracy or even complete interruption of output, directly impairing their ability to communicate.

For individuals with severe speech impairment, such disruptions have immediate and significant consequences for autonomy and daily functioning, as communication may be entirely dependent on the BCI system. Clinical management would therefore require immediate fallback to alternative communication methods, such as eye-tracking systems, switch-based interfaces, or caregiver-assisted communication, to maintain continuity of interaction. For implantable systems, additional considerations include the need for surgical revision, device replacement, or controlled deactivation in cases of hardware degradation, infection, or long-term signal instability. At present, most clinical speech BCI studies are conducted in highly supervised research environments, involving a small number of participants with continuous technical support, frequent recalibration, and direct researcher oversight<sup>1,2,3</sup>. However, standardized clinical protocols for failure detection, response, long-term maintenance, and emergency communication backup have not yet been systematically established in the literature. Ensuring patient safety, system reliability, and uninterrupted access to communication will therefore be a critical requirement for the transition of speech BCIs from experimental systems to dependable clinical technologies. Importantly, these ethical and clinical considerations are derived from studies involving implanted systems in target patient populations and should be clearly distinguished from foundational speech decoding research conducted in non-clinical settings, which does not involve comparable surgical or long-term care concerns.

### **Cognitive Privacy, Data Governance, and Consent Frameworks**

Beyond the risks associated with surgical implantation, speech BCIs introduce unique ethical challenges related to the collection, processing, and use of neural data within current decoding systems. In clinical speech BCI studies involving participants with severe paralysis such as those with amyotrophic lateral sclerosis (ALS) or anarthria, intracortical and electrocorticography (ECoG)-based systems acquire continuous neural recordings from speech-related cortical regions and process them through machine learning models to reconstruct linguistic output<sup>2,3,17</sup>. These recordings consist of high-dimensional time-series signals that are not inherently limited to task-relevant activity and may therefore contain incidental or non-task-related neural information. While similar neural decoding principles are applied in foundational speech neuroscience studies conducted in non-paralyzed participants, the ethical considerations discussed here arise specifically in clinical contexts, where such systems are used for direct communication in individuals with severe speech impairment.

Within these clinical settings, current studies typically constrain decoding through predefined experimental paradigms,

---

such as prompted phrases or attempted speech tasks. Nevertheless, the underlying neural data streams are broader in scope, creating a theoretical risk—particularly in real-time communication environments that decoded outputs could extend beyond intended communication if these constraints are relaxed<sup>6</sup>. This raises a central issue of ownership and control of neural data. In present clinical research environments, neural recordings are stored, processed, and analyzed within institutionally governed systems under formal ethics approval, with access restricted to study teams<sup>2,3</sup>. However, standardized frameworks for long-term data ownership, particularly in scenarios involving cloud-based processing, multi-institutional sharing, or incorporation of neural datasets into commercial algorithm development have not yet been established. Given that neural data may encode aspects of internal cognitive processes, there is a strong ethical argument for models in which patients retain primary control over how their data are stored, accessed, and reused, although such models have not yet been systematically implemented or evaluated in clinical BCI contexts.

A second key requirement is the restriction of decoding to intentional communication signals. High-performance clinical speech BCI systems currently achieve low error rates in part because they operate under tightly controlled paradigms, such as cued attempted speech or structured phrase generation<sup>2,17</sup>. These paradigms inherently constrain decoding to task-relevant neural activity. However, as speech BCIs move toward continuous, real-world clinical deployment, particularly when exploring domains like inner (imagined) speech studied in foundational or early-stage research, distinguishing intentional communication from spontaneous neural activity becomes far more complex<sup>8</sup>. Addressing this challenge will likely require explicit control mechanisms, such as user-initiated activation signals, neural gating thresholds that suppress low-confidence outputs, or confirmation-based interfaces that validate decoded content prior to transmission. Currently, such safeguards are neither standardized nor widely evaluated in clinical studies.

Data security and system integrity represent an additional layer of concern. Speech BCIs rely on real-time acquisition, transmission, and processing of neural signals through external hardware and computational systems. At present, these processes occur within controlled clinical or laboratory environments under continuous technical oversight<sup>17</sup>). As systems transition toward portable or home-based clinical use, risks related to unauthorized data access, signal interception, or system manipulation will become increasingly relevant. Ensuring secure implementation will require end-to-end encryption of neural data streams, authenticated access controls, and auditable system logs, though such measures are not consistently described in existing primary research literature.

Finally, consent frameworks must account for the longitu-

dinal and adaptive nature of clinical speech BCI use. Unlike conventional clinical interventions, these systems involve continuous data collection, repeated model retraining, and potential secondary use of neural recordings for algorithm improvement. Current studies obtain informed consent within structured research protocols, but do not address long-term scenarios in which neural data are stored indefinitely, shared across institutions, or incorporated into evolving machine learning systems<sup>2,3</sup>. This highlights the need for dynamic consent models that allow patients to update preferences over time, including the ability to withdraw participation, request deletion of stored data, or restrict specific forms of data reuse. Such considerations are particularly critical in progressive neurological conditions, where decision-making capacity and communication ability may change over time. Taken together, these issues indicate that ethical safeguards for speech BCIs must extend beyond general principles to include concrete technical, clinical, and regulatory mechanisms governing data use, system behavior, and patient autonomy. At present, these frameworks remain underdeveloped, particularly in the context of real-world clinical deployment for individuals with severe paralysis, and are not systematically addressed in existing primary research studies representing a critical barrier to responsible and scalable clinical translation.

### **Global Accessibility, Affordability, and Health-System Barriers**

A critical but often underemphasized barrier to clinical reality, particularly for speech BCIs developed and evaluated in small clinical cohorts of individuals with severe paralysis, is their global feasibility. Current high-performance clinical speech BCI systems, as demonstrated in small-scale studies involving participants with severe speech impairment, rely on expensive hardware, advanced neurosurgical expertise, prolonged training protocols, and continuous technical support<sup>2,3,17</sup>. Consequently, access is currently restricted to a small number of well-funded research centers, predominantly located in high-income countries. These observations are derived from highly specialized research environments rather than large-scale healthcare implementation. In low- and middle-income regions, additional barriers include limited neurosurgical infrastructure, shortages of trained clinicians and engineers, and competing public health priorities. Even if technical performance continues to improve, the absence of scalable training programs for surgeons, rehabilitation specialists, and clinical technicians could significantly limit scalable real-world clinical deployment. Achieving true clinical reality will therefore require parallel innovation beyond the laboratory. Cost-reduction strategies, modular or semi-invasive system designs, standardized training pipelines, and integration into existing healthcare frameworks will be essential. Without deliberate

---

efforts to address affordability and workforce development, there is a substantial risk that speech BCIs, if not deliberately designed for scalability and accessibility, could exacerbate existing global health inequalities rather than alleviate them.

## Limitations of Current Research

### *Small Sample Sizes*

A consistent limitation across clinical speech BCI studies involving participants with severe paralysis or anarthria is the extremely small number of participants, often involving fewer than five individuals in single-case or small cohort designs. While this is understandable given the invasive nature of the technology and the rarity of eligible participants, such small cohorts limit statistical generalizability<sup>2,3,17</sup> with similar small-sample constraints also observed in related foundational intracortical decoding studies<sup>5</sup>. Individual differences in cortical organization, disease progression, and learning capacity may substantially influence outcomes, complicating predictions for broader patient populations. This limitation is particularly significant when evaluating claims related to clinical scalability and reproducibility across diverse neurological conditions. **Limited Long-Term Data**

Although several clinical speech BCI studies in participants with severe speech impairment demonstrate stable performance over periods of months, most trials remain relatively short<sup>2,3,17</sup>. Long-term challenges such as signal degradation, device longevity, user fatigue, and psychological adaptation over multiple years remain insufficiently explored. Without extended longitudinal data, it remains premature to fully assess the long-term durability, reliability, and real-world practicality of speech BCIs as continuous clinical communication systems.

The evaluation of psychosocial outcomes and safety profiles in clinical speech BCI studies is further constrained by the absence of standardized and consistently reported measures across studies, including variability in outcome definitions, reporting formats, follow-up durations, and methods of adverse-event documentation. As a result, direct comparison and quantitative aggregation of these outcomes were not methodologically feasible. This limitation reflects both small sample sizes and the early-stage, experimental nature of clinical speech BCI research. Accordingly, the findings related to patient experience and safety should be interpreted as qualitative, study-level trends rather than systematically derived or statistically validated evidence.

### **Future Directions**

Future progress toward clinically viable speech BCIs for individuals with severe paralysis will depend on continued advances in neural recording technology, decoding algorithms,

and speech synthesis models, as informed by both clinical studies and foundational neural decoding research<sup>18,38</sup>. Improvements in electrode longevity, wireless data transmission, and adaptive machine-learning systems could reduce calibration demands and improve robustness in non-laboratory environments. However, most of these improvements have not yet been validated in long-term clinical use in individuals with severe speech impairment. Hybrid or semi-invasive approaches may help balance performance with safety and accessibility. Inner (imagined) speech decoding represents a potentially transformative research direction, although it remains largely at a foundational or early-stage experimental level and has not yet been validated in clinical populations with severe speech impairment<sup>8</sup>. This approach will necessitate robust and explicitly defined ethical safeguards to ensure that only intentional communication is decoded and that cognitive privacy is preserved.

Crucially, the transition from experimental demonstrations to clinically reliable communication systems will require large-scale, multi-center clinical trials. Such trials should include diverse patient populations, standardized outcome measures, and long-term follow-up, integrating technical, clinical, psychological, and social metrics. Such efforts will be essential to determine whether performance observed in small, highly controlled studies can be reproduced across diverse patient populations and sustained over clinically meaningful timeframes.

### **Closing Perspective: Are We Truly Near Clinical Reality?**

The evidence reviewed in this report suggests that speech-restoring BCIs are approaching but have not yet fully achieved clinical reality. Technologically, invasive speech BCIs have demonstrated the ability to restore meaningful, expressive communication in a very small number of individuals with severe paralysis under highly controlled experimental conditions, rather than across broader clinical populations<sup>2,3,17</sup>. Within the target clinical population, evidence remains limited to a small number of participants, typically between one and five individuals studied under intensive researcher-supported experimental conditions, in whom qualitative reports describe perceived gains in autonomy, identity, and quality of life<sup>2,3,17</sup>. However, these observations are derived from highly selected participants receiving intensive technical and clinical support and are not yet generalizable to broader patient populations. Ethically and practically, however, challenges related to invasiveness, cost, scalability, global accessibility, and long-term reliability remain unresolved. Speech BCIs can therefore be considered clinically feasible in highly controlled experimental contexts for select individuals, but not yet established as generalizable or routinely deployable clinical interventions. Whether these remarkable experimental achievements evolve

---

into accessible, equitable, and sustainable clinical tools will depend on continued interdisciplinary collaboration ensuring that the future of speech restoration addresses not only technological possibility, but the human experience of communication itself.

When evaluated against the criteria for clinical reality defined in this review, current speech brain–computer interfaces (BCIs) demonstrate substantial technological progress based on evidence from both small-scale clinical studies in individuals with paralysis and foundational neural decoding research, but do not yet satisfy all conditions required for routine clinical deployment. Each benchmark communication speed, decoding accuracy, long-term stability, reproducibility across participants, and clinical accessibility represents a practical requirement that determines whether a system can function as a reliable therapeutic communication interface rather than a controlled experimental prototype. The first criterion concerns communication speed, defined in this review as the ability of a speech BCI to generate intelligible linguistic output at rates approaching 50–70 words per minute, which corresponds approximately to the lower range of natural conversational speech. This benchmark is used because effective spoken interaction depends on rapid turn-taking and the ability to construct sentences without extended pauses between words. Communication technologies that operate substantially below this range force users to communicate through slow, segmented message construction that disrupts conversational flow. Conventional assistive communication systems such as eye-tracking keyboards and switch-based spellers typically produce output at approximately 5–20 words per minute, because each character must be selected individually. In contrast, speech BCIs attempt to decode neural activity associated with speech planning and articulation directly from cortical motor regions, allowing multiple phonemes or entire words to be generated from continuous neural signals rather than sequential character selection. Recent invasive electrocorticography-based systems have demonstrated output rates approaching 60 words per minute in individual participants with severe paralysis under structured experimental conditions, achieved through high-density cortical recordings combined with machine-learning decoders that translate neural activity patterns into text or synthesized speech. These results indicate that the communication throughput required for conversational interaction may be achievable in controlled experimental settings, although current evidence is limited to small clinical samples and does not yet establish consistent real-world performance. However, because such performance has primarily been demonstrated in controlled laboratory environments and limited conversational contexts, the communication speed criterion can reasonably be considered partially satisfied but not yet fully validated in everyday clinical use.

The second criterion concerns decoding accuracy, opera-

tionalized here as a word error rate below approximately 20–25% for open-vocabulary speech decoding. This threshold reflects the level at which listeners can typically reconstruct intended meaning without extensive manual correction or repeated clarification. In natural language comprehension, communication becomes increasingly difficult when error rates exceed roughly one incorrect word in four, because contextual cues are no longer sufficient to reliably infer missing information. Early speech BCI systems were largely restricted to small predefined vocabularies or isolated phoneme classification tasks, which limited their usefulness for spontaneous communication. Recent advances in neural decoding algorithms, many of which are informed by foundational studies conducted in non-clinical populations or controlled experimental settings, have improved performance by allowing systems to interpret neural activity patterns in the context of likely linguistic sequences. As a result, several recent demonstrations have reported word error rates approaching or slightly below the 25% threshold when decoding structured sentences or moderately constrained vocabularies. These improvements suggest that neural representations of speech articulation can be translated into intelligible linguistic output under controlled conditions, although clinical validation in target patient populations remains limited. Nevertheless, error rates remain higher when decoding unconstrained spontaneous speech, and performance can vary across recording sessions and participants. Consequently, although current systems demonstrate substantial progress toward clinically useful accuracy, the decoding reliability required for unrestricted real-world communication has not yet been consistently achieved.

The third criterion concerns long-term stability, defined as the ability of a speech BCI to maintain consistent decoding performance over extended periods of at least 6–12 months without frequent recalibration or signal loss. This requirement reflects the practical needs of individuals with severe speech paralysis, who must rely on communication technologies continuously in daily life. Neural interfaces can experience changes in signal quality over time due to biological responses at the electrode–tissue interface, gradual shifts in neural activity patterns, or hardware degradation. Although several studies have demonstrated stable neural recordings over periods of months, systematic evidence demonstrating consistent high-accuracy speech decoding maintained across year-long timescales remains limited. Many experimental systems still require periodic retraining of decoding algorithms to compensate for changes in signal characteristics. While recalibration procedures may be manageable within research environments, they could impose significant burdens on patients if required frequently during everyday use. For this reason, the long-term stability criterion cannot yet be considered fully satisfied, and additional longitudinal studies will be necessary to establish whether neural speech decoding performance can

---

remain reliable across the multi-year timescales required for clinical assistive technologies.

A fourth criterion involves reproducibility across participants and research environments, which is essential for determining whether the technology can be generalized beyond isolated proof-of-concept demonstrations. Clinical interventions must demonstrate consistent performance across individuals with different neurological conditions, cortical anatomies, and patterns of functional impairment. Many of the most advanced demonstrations of speech BCI performance in the target clinical population have been conducted with very small participant cohorts, sometimes involving only one or two individuals implanted with experimental neural recording devices. High-performance speech decoding in the target clinical population currently depends primarily on invasive neural recording technologies, such as electrocorticography arrays or intracortical microelectrode implants, which require neurosurgical procedures for implantation. Differences in electrode placement, disease progression, and cortical reorganization may substantially influence decoding performance. As a result, large-scale multi-participant validation studies remain necessary to determine whether speech BCIs can provide reliable communication across diverse users, meaning that the reproducibility criterion is currently only partially satisfied.

The final criterion concerns clinical accessibility and translational feasibility, referring to whether a speech BCI system can realistically be implemented within healthcare systems outside specialized research laboratories. High-performance speech decoding currently depends primarily on invasive neural recording technologies, such as electrocorticography arrays or intracortical microelectrode implants, which require neurosurgical procedures for implantation. Although these devices provide the spatial and temporal resolution necessary to capture detailed motor speech signals, their use introduces considerations related to surgical risk, long-term device durability, and regulatory approval processes governing implantable neurotechnology. In addition, many existing experimental systems require specialized computational infrastructure and continuous technical oversight by research teams. For widespread clinical adoption, speech BCIs would need to operate reliably within standard clinical environments and eventually within home settings with minimal technical intervention. Because these translational challenges remain unresolved, clinical accessibility currently represents one of the most significant barriers to the widespread deployment of speech BCI technologies.

Taken together, the available evidence from both target clinical studies and foundational neural decoding research indicates that contemporary speech BCIs have achieved substantial progress toward the functional goals required for neural speech restoration, particularly in terms of communication speed and decoding capability. However, the remaining chal-

lenges in long-term stability, large-scale reproducibility, and practical clinical implementation indicate that the technology has not yet reached the stage defined in this review as clinical reality. Instead, current systems should be regarded as advanced experimental platforms that demonstrate the feasibility of direct neural speech communication but still require further engineering development, longitudinal validation, and translational testing before they can become routine clinical communication tools for individuals with severe speech paralysis. When evaluated directly against the criteria established in this review, current speech BCIs partially satisfy the requirements for clinical reality, with the strongest evidence derived from small-scale clinical studies in individuals with paralysis and supported by broader foundational decoding research, but they do not fully meet all criteria. Specifically, the communication speed criterion is largely met, because several invasive systems have demonstrated output rates approaching conversational speech levels under controlled conditions. The decoding accuracy criterion is approaching fulfillment but remains only partially satisfied, since reported word error rates in some systems fall near the proposed intelligibility threshold yet still vary across contexts and participants. In contrast, the long-term stability criterion is not yet satisfied, because consistent speech decoding performance maintained over multi-year timescales has not been systematically demonstrated. Similarly, the reproducibility criterion remains only partially satisfied, as most high-performance demonstrations have been limited to small participant samples and have not yet been replicated across large patient populations or multiple independent clinical research centers. Finally, the clinical accessibility criterion is not currently satisfied, because the most effective speech decoding systems rely on invasive neural implants and complex research infrastructure that have not yet been translated into widely deployable clinical devices. For these reasons, although speech BCIs have demonstrated clear technological viability and strong therapeutic potential, the available evidence indicates that they cannot yet be classified as fully realized clinical communication technologies. Rather, they should currently be understood as advanced translational neurotechnology systems that are approaching clinical feasibility but have not yet reached full clinical reality, as defined by the performance, reliability, reproducibility, and accessibility standards outlined in this review.

## Conclusion

This review examined whether speech-restoring brain-computer interfaces (BCIs) meet the criteria for clinical reality by synthesizing evidence from both target clinical populations and foundational neural decoding studies while evaluating decoding accuracy, communication speed, latency, long-term stability, reproducibility, and translational feasibility.

ity. The evidence indicates that substantial technical progress has been achieved, particularly in invasive systems using intracortical microelectrode arrays and electrocorticography (ECoG). Under controlled experimental conditions, single-participant studies (typically  $n = 1-5$ ) have demonstrated word error rates of approximately 9% in small-vocabulary, closed-set paradigms and approximately 15–25% in larger, open-vocabulary decoding tasks, alongside communication speeds approaching , 60 words per minute and latencies below 100–200 milliseconds. These results confirm that neural signals from speech-related cortical regions can be decoded into intelligible linguistic output at rates approaching the lower bound of conversational speech in controlled experimental conditions, with clinical demonstrations currently limited to a small number of participants with severe paralysis. However, these performance levels are highly contingent on specific experimental conditions, including structured task design (e.g., cued or prompted speech), extensive participant-specific training, integration of probabilistic language models, and continuous technical supervision. As such, they represent optimized, upper-bound performance in controlled laboratory or clinical environments rather than generalized real-world outcomes. Moreover, the current evidence base is limited by extremely small sample sizes, short follow-up durations (typically weeks to months), and a lack of standardized evaluation protocols, making it difficult to assess variability across users, long-term reliability, or robustness in unconstrained communication settings.

When evaluated against the criteria defined in this review, speech BCIs partially satisfy the requirements for clinical reality. Communication speed and latency benchmarks have been achieved in select invasive systems under controlled conditions. Decoding accuracy is approaching clinically usable thresholds but remains variable, particularly in open-vocabulary and spontaneous communication contexts. In contrast, long-term stability beyond several months, reproducibility across diverse patient populations, and scalability within routine healthcare systems remain insufficiently demonstrated. Non-invasive systems, while safer and more accessible, continue to exhibit lower communication speeds (, 15–20 words per minute) and higher, more variable error rates, limiting their current clinical utility for continuous speech restoration. Emerging paradigms such as neural-to-speech synthesis and inner (imagined) speech decoding further expand the potential scope of speech BCIs, but remain at early or proof-of-concept stages, often evaluated in non-target populations or under constrained and partially offline conditions. These approaches therefore represent long-term research directions rather than near-term clinical solutions.

Taken together, the current evidence supports the conclusion that speech BCIs have achieved demonstrated clinical feasibility in highly controlled, small-sample studies within

the target population, supported by broader foundational research, but have not yet reached reproducible, scalable, and long-term clinical reality as defined by consistent performance across users, sustained operation over extended timeframes, and integration into standard healthcare environments. Bridging this gap will require not only further improvements in decoding performance, but also systematic validation across larger cohorts, development of standardized clinical protocols, and resolution of key translational challenges, including device reliability, failure management, data governance, and accessibility. Future progress will depend on multi-center studies with larger and more diverse participant populations, longitudinal evaluation extending beyond months to years, and the integration of technical, clinical, and ethical frameworks into cohesive deployment models. Only through addressing these interconnected requirements can BCIs transition from high-performance experimental systems to reliable, widely accessible clinical communication technologies.

## Acknowledgments

The author would like to acknowledge the researchers, clinicians, and participants whose work made this review possible. In particular, gratitude is extended to individuals living with severe paralysis who have volunteered to participate in brain–computer interface research, often undergoing demanding experimental protocols to advance scientific understanding and future clinical care. Their contributions are central to progress in speech-restoring neurotechnology. The author also acknowledges the scientific community whose open dissemination of peer-reviewed research, datasets, and methodological frameworks enabled this synthesis. No external funding was received for this work. The author declares no conflicts of interest.

## References

- 1 F. R. Willett, et al., A high-performance speech neuroprosthesis. *Nature*, 620, 1031–1036 (2023). doi:10.1038/s41586-023-06377-x.
- 2 S. L. Metzger, et al., A high-performance neuroprosthesis for speech decoding and avatar control. *Nature*, 620, 1037–1046 (2023). doi:10.1038/s41586-023-06443-4.
- 3 D. A. Moses, et al., Neuroprosthesis for decoding speech in a paralyzed person with anarthria. *New England Journal of Medicine*, 385, 217–227 (2021). doi:10.1056/NEJMoa2027540.
- 4 G. K. Anumanchipalli, et al., Speech synthesis from neural decoding of spoken sentences. *Nature*, 568, 493–498 (2019). doi:10.1038/s41586-019-1119-1.
- 5 C. Pandarinath, et al., High performance communication by people with paralysis using an intracortical brain–computer interface. *eLife* (2017). doi:10.7554/eLife.18554.

- 6 J. G. Makin, et al., Machine translation of cortical activity to text with an encoder–decoder framework. *Nature Neuroscience*, 23, 575–582 (2020). doi:10.1038/s41593-020-0608-8.
- 7 C. Herff, et al., Brain-to-text: decoding spoken phrases from phone representations in the brain. *Frontiers in Neuroscience*, 9, 217 (2015). doi:10.3389/fnins.2015.00217.
- 8 J. T. Panachakel and A. G. Ramakrishnan, Decoding covert speech from EEG: a comprehensive review. *Frontiers in Neuroscience*, 15, 642251 (2021). doi:10.3389/fnins.2021.642251.
- 9 S. D. Stavisky, et al., Neural ensemble dynamics in dorsal motor cortex during speech in people with paralysis. *eLife*, 8, e46015 (2019). doi:10.7554/eLife.46015.
- 10 G. H. Wilson, et al., Decoding spoken English from intracortical electrode arrays in dorsal precentral gyrus. *Journal of Neural Engineering*, 17, 066007 (2020). doi:10.1088/1741-2552/abbfef.
- 11 A. Tankus, et al., Structured auditory representations in the human superior temporal gyrus. *Nature*, 485, 233–236 (2012). doi:10.1038/nature11023.
- 12 K. E. Bouchard, N. Mesgarani, K. Johnson, E. F. Chang, Functional organization of human sensorimotor cortex for speech articulation. *Nature* (2013). doi:10.1038/nature11911.
- 13 X. Pei, D. L. Barbour, E. C. Leuthardt, G. Schalk, Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans. *Journal of Neural Engineering* (2011). doi:10.1088/1741-2560/8/4/046028.
- 14 F. H. Guenther, et al., A wireless brain-machine interface for real-time speech synthesis. *PLOS ONE*, 4, e8218 (2009). doi:10.1371/journal.pone.0008218.
- 15 F. Bocquelet, et al., Real-time control of a formant-based speech synthesizer using a low-latency intracortical brain–computer interface. *Journal of Neural Engineering*, 13, 026005 (2016). doi:10.1088/1741-2560/13/2/026005.
- 16 N. S. Card, et al., An accurate and rapidly calibrating speech neuroprosthesis. *New England Journal of Medicine*, 391, 609–618 (2024). doi:10.1056/NEJMoa2314132.
- 17 F. R. Willett, D. T. Avansino, L. R. Hochberg, J. M. Henderson, K. V. Shenoy, High-performance brain-to-text communication via handwriting. *Nature* (2021). doi:10.1038/s41586-021-03506-2.
- 18 M. Angrick, et al., Speech synthesis from ECoG using densely connected 3D convolutional neural networks. *Journal of Neural Engineering* (2019). doi:10.1088/1741-2552/ab0c59.
- 19 H. Akbari, et al., Towards reconstructing intelligible speech from the human auditory cortex. *Scientific Reports* (2019). doi:10.1038/s41598-018-37359-z.
- 20 N. Mesgarani, et al., Phonetic feature encoding in human superior temporal gyrus. *Science* (2014). doi:10.1126/science.1245994.
- 21 B. Khalighinejad, et al., The sequence of neural events in speech production. *Nature Communications* (2017). doi:10.1038/ncomms14685.
- 22 A. Craik, Y. He, J. L. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: A review. *Journal of Neural Engineering* (2019). doi:10.1088/1741-2552/ab0ab5.
- 23 Y. Roy, et al., Deep learning-based electroencephalography analysis: A systematic review. *Journal of Neuroscience Methods* (2019). doi:10.1016/j.jneumeth.2019.108441.
- 24 J. Yuan, M. Liberman, C. Cieri, Towards an integrated understanding of speaking rate in conversation. *Interspeech* (2006). doi:10.21437/Interspeech.2006-204.
- 25 N. Rahman, et al., Advances in brain–computer interface for decoding speech imagery from EEG signals: A systematic review. (2024). doi:10.1007/s11571-024-10167-0.
- 26 D. Lopez-Bernal, et al., A state-of-the-art review of EEG-based imagined speech decoding. *Frontiers in Human Neuroscience* (2022). doi:10.3389/fnhum.2022.867281.
- 27 P. V. Dhole, et al., EEG-based brain–computer interfaces for imagined speech classification. (2024). doi:10.53555/kuey.v30i5.3894.
- 28 M. V. Hareesh and B. S. Begum, Identification of brain states from EEG signals for imagined speech BCIs. *Behavioural Brain Research* (2025). doi:10.1016/j.bbr.2024.115295.
- 29 K. Bhadra, et al., Learning to operate an imagined speech brain–computer interface. *Communications Biology* (2025). doi:10.1038/s42003-025-07464-7.
- 30 C. H. Nguyen, et al., EEG-based imagined speech classification using deep learning. *Neurocomputing* (2020). doi:10.1016/j.neucom.2020.02.101.
- 31 C. Cooney, R. Folli, D. Coyle, Neurolinguistics research advancing development of a direct-speech brain–computer interface. *iScience* (2018). doi:10.1016/j.isci.2018.09.016.
- 32 C. A. Chestek, et al., Long-term stability of neural prosthetic control signals from silicon cortical arrays. *Journal of Neural Engineering* (2011). doi:10.1088/1741-2560/8/4/045005.
- 33 J. A. Perge, et al., Intra-day signal instabilities affect decoding performance in intracortical neural interface systems. *Journal of Neural Engineering* (2013). doi:10.1088/1741-2560/10/3/036004.
- 34 L. R. Hochberg, et al., Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* (2012). doi:10.1038/nature11076.
- 35 M. Enca and R. Andorno, Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy* (2017). doi:10.1186/s40504-017-0050-1.
- 36 R. Yuste, et al., Four ethical priorities for neurotechnologies and AI. *Nature* (2017). doi:10.1038/551159a.
- 37 E. M. Mugler, et al., Direct classification of all American English phonemes using signals from functional speech motor cortex. *Journal of Neural Engineering* (2014). doi:10.1088/1741-2560/11/3/035015.
- 38 J. Chartier, G. K. Anumanchipalli, K. Johnson, E. F. Chang, Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex. *Neuron* (2018). doi:10.1016/j.neuron.2018.04.031.
- 39 D. Dash, P. Ferrari, J. Wang, Decoding imagined and spoken phrases from non-invasive neural (MEG) signals. *Frontiers in Neuroscience*, 14, 290 (2020). doi:10.3389/fnins.2020.00290.
- 40 S. Martin, et al., Word pair classification during imagined speech using direct brain recordings. *Scientific Reports* (2016). doi:10.1038/srep25803.
- 41 T. Proix, et al., Imagined speech can be decoded from intracranial neural activity. *Nature Communications* (2022). doi:10.1038/s41467-021-27725-3.