The Clinical Horizon for Non-Invasive Speech Neuroprostheses: A State-of-the-Art Analysis for Acute Care Applications

Executive Summary

The restoration of communication for individuals rendered non-verbal by neurological injury or disease represents one of the most profound challenges in modern medicine. Brain-Computer Interfaces (BCIs) that can decode neural signals into intelligible speech offer a transformative solution, particularly for patients in acute care settings like the Intensive Care Unit (ICU) who have lost the ability to express their needs, pain, or wishes. While invasive BCIs, requiring neurosurgical implantation, have demonstrated remarkable progress, achieving near-conversational decoding speeds, their associated risks limit their widespread applicability. The ultimate goal for broad clinical deployment, especially in transient or acute conditions, is a non-invasive system that is safe, rapidly deployable, and effective.

This report provides an exhaustive analysis of the current state-of-the-art and future trajectory for developing such a clinically viable, non-invasive speech BCI. The central finding is that a significant performance chasm exists between invasive and non-invasive modalities, rooted in the fundamental physics of signal acquisition. Non-invasive Electroencephalography (EEG), the most promising modality for this application, is plagued by a low signal-to-noise ratio and poor spatial resolution, which have historically limited its use to classifying small, discrete vocabularies.

However, the field is at an inflection point, driven by convergent advancements across multiple domains. First, hardware innovations in dry-electrode EEG systems, integrated amplifiers, and wearable, high-density arrays are beginning to overcome the long-standing barriers of patient comfort and setup time, making long-term, continuous monitoring in an ICU setting feasible. Second, a paradigm shift in software from traditional feature engineering to end-to-end deep learning models, particularly those based on Transformer architectures, has unlocked new potential for decoding

complex neural patterns directly from raw data. Third, the critical challenge of lengthy, patient-specific calibration is being addressed by emerging techniques in few-shot and zero-shot learning, which promise to create models that can be adapted to new users with minimal data.

Despite this progress, formidable translational hurdles remain. The unique physiological and cognitive states of ICU patients, including fatigue, fluctuating consciousness, and the effects of medication, present immense challenges to BCI stability and performance. Furthermore, the chaotic ICU environment is electrically hostile, demanding highly robust artifact suppression algorithms. Overcoming these barriers requires a pivot to a user-centered design philosophy, focusing not only on technical benchmarks like Word Error Rate (WER) and Words Per Minute (WPM) but also on usability, cognitive load, and the design of intuitive interfaces for both patients and caregivers.

Finally, the report underscores the neuroethical imperative that must guide this research. As the technology inches closer to decoding the contents of thought, profound questions of mental privacy, informed consent for non-verbal individuals, responsible management of expectations, and the mitigation of decoding errors come to the forefront. The successful translation of non-invasive speech BCIs from the laboratory to the ICU bedside will depend not only on surmounting the immense technical challenges but also on navigating this complex ethical landscape with foresight and care.

Section 1: The Modality Chasm: Bridging the Performance Gap in Neural Signal Acquisition

The central challenge in developing a clinically viable non-invasive speech BCI is the immense disparity in signal quality, and consequently decoding performance, between methods that record from the scalp and those that record directly from the brain. This section dissects the physical and practical underpinnings of this performance gap, establishing the fundamental limitations of non-invasive techniques and exploring the hardware innovations that aim to narrow this chasm. The performance of an invasive system is largely an engineering problem centered on decoding a high-fidelity signal, whereas the performance of a non-invasive system remains a fundamental physics problem of first extracting a usable signal from overwhelming noise.

1.1. The EEG Conundrum: Theoretical Limits and Practical Realities of Scalp-Based Speech Decoding

Electroencephalography (EEG) stands as the most promising modality for a deployable, non-invasive speech BCI due to its high temporal resolution, portability, and low cost.[1] However, its clinical application for complex tasks like speech synthesis is severely constrained by fundamental physical limitations.

The primary obstacle is an inherently low signal-to-noise ratio (SNR).[1] The electrical signals generated by cortical neurons are weak, and they must traverse the skull, dura, and scalp tissues before reaching the recording electrodes. These layers act as a spatial low-pass filter, a phenomenon known as volume conduction, which attenuates the signal and smears it across the scalp.[1] The neural signatures associated with

imagined speech, which are subtler than those for overt action, are particularly vulnerable to being lost in the noise of background brain activity and non-neural artifacts.[1]

Compounding the SNR problem is EEG's poor spatial resolution. While it can capture neural dynamics on a millisecond timescale, its ability to localize the source of those signals is on the order of centimeters.[3] This makes it profoundly difficult to disentangle the activity of adjacent but functionally distinct cortical areas that are co-activated during speech processing.[1] The fine-grained spatiotemporal patterns that encode phonetic and articulatory information, which are accessible to invasive methods, become blurred and overlapping at the scalp.

These physical limitations have direct practical consequences on decoding performance. For decades, EEG-based speech decoding has struggled to move beyond simple classification tasks. While binary classification (e.g., discriminating between two words) can achieve accuracies exceeding 80% in controlled settings, performance drops precipitously as the number of classes increases.[1] Multi-class classification, essential for any practical communication system, often yields accuracies that are only slightly above chance level.[7] This has historically confined EEG-based systems to extremely limited vocabularies, such as a handful of vowels or basic command words like "up" or "down," rendering them insufficient for meaningful conversation.[1]

1.2. Quantifying the Divide: A Comparative Analysis of Invasive and Non-Invasive BCI Performance

The performance gap between invasive and non-invasive methods is not incremental; it is a categorical divide that separates systems demonstrating clinical feasibility from those that remain largely experimental.

Invasive technologies, such as electrocorticography (ECoG) grids placed on the cortical surface and intracortical microelectrode arrays that penetrate the cortex, bypass the signal-distorting skull entirely.[8] This direct access to neural tissue yields signals with vastly superior SNR and spatial resolution. High-density micro-electrocorticography (μECoG) arrays, for instance, can capture signals with a 48% higher SNR compared to standard clinical intracranial EEG (IEEG), leading to a 35% improvement in speech decoding.[10]

This superior signal fidelity translates into remarkable decoding performance. State-of-the-art invasive BCIs have demonstrated the ability to decode attempted speech into text in real-time at speeds approaching natural conversation. Published results report decoding rates of 62 to 78 words per minute (WPM).[11] The associated accuracy, measured by Word Error Rate (WER), is impressively low, achieving 9.1% on a 50-word vocabulary and a 23.8% on a large, conversational vocabulary of 125,000 words.[14] These systems represent a functional proof-of-concept for restoring communication.

In stark contrast, the performance of non-invasive systems lags far behind.

- **EEG:** As discussed, performance is a major hurdle. Even recent, highly optimized studies using deep learning struggle. One landmark study achieved a 48% top-1 accuracy on a 512-phrase classification task, but this required an unprecedented 175 hours of training data from a single participant, highlighting the immense data requirements to overcome the poor SNR.[17] Most research remains focused on classifying very small, discrete vocabularies with accuracies in the 70-90% range, which does not translate to open-vocabulary synthesis.[5]
- **Magnetoencephalography (MEG):** MEG measures the magnetic fields produced by neural currents. Because these fields are not distorted by the skull, MEG offers

better signal quality and spatial resolution than EEG.[20] Studies have shown that achieving significant speech-brain tracking requires approximately three times less recording time with MEG than with EEG.[20] However, MEG systems are extraordinarily expensive and require large, magnetically shielded rooms, making them completely impractical for deployment in an ICU or any other clinical setting outside of a specialized research facility.[3]
- **Functional Near-Infrared Spectroscopy (fNIRS):** fNIRS is an optical technique that measures changes in blood oxygenation related to neural activity. It is portable, robust to motion artifacts, and has better spatial resolution than EEG.[23] However, it measures the slow hemodynamic response, which lags behind neural activity by several seconds.[17] This poor temporal resolution makes fNIRS fundamentally unsuitable for the real-time decoding of continuous, rapid speech events.[17] Its primary potential lies in hybrid systems that combine fNIRS with EEG to leverage the spatial information of the former and the temporal precision of the latter.[25]

The fundamental trade-off is thus clear: invasive methods provide high performance at the cost of significant surgical risk, patient burden, and limited accessibility, while non-invasive methods are safe and scalable but are hamstrung by poor signal quality that severely limits their decoding capabilities.[17]

Table 1: Comparative Analysis of Neural Acquisition Modalities

Modality	Invasiveness	Temporal Resolution	Spatial Resolution	Signal-to-Noise Ratio (SNR)	Key Limitations for ICU Use	State-of-the-Art Speech Performance
Intracortical Arrays	High (penetrating)	<1 ms (spikes)	<0.1 mm	Very High	Surgical risk, infection, signal degradation over time	62-78 WPM; 9-24% WER [11]
ECoG / μECoG	High (surface)	~1 ms	~1–4 mm	High	Surgical risk,	35% decoding

					limited coverage	improvement over IEEG [10]
High-Density EEG	Non-invasive	~1 ms	~25 mm2	Low	Volume conduction, artifacts, setup time (wet)	~48% accuracy on 512-phrase task [17]
MEG	Non-invasive	~1 ms	~1 mm	Moderate	Non-portable, extremely expensive, requires shielded room	Better than EEG, but not clinically deployed [22]
fNIRS	Non-invasive	Seconds	~10–20 mm	Moderate	Slow hemodynamic response, limited depth	Not suitable for real-time continuous speech [17]

1.3. The Hardware Frontier: Engineering the Clinically Viable Non-Invasive Interface

Given the profound limitations of non-invasive signals, progress toward a clinically viable system is disproportionately dependent on upstream hardware innovations that can improve the quality of the raw signal at its source. The most significant advancements are in EEG electrode and system design, aimed at making the technology practical for rapid, long-term use in challenging environments like the ICU.

The evolution from traditional "wet" electrodes to modern "dry" systems is a critical step. Wet electrodes require the application of a conductive gel or paste to reduce skin impedance, a process that is time-consuming, requires a trained technician, is uncomfortable for the patient, and is unsuitable for long-term monitoring as the gel dries out and degrades signal quality.[31] These factors make wet EEG impractical for routine ICU use.

Dry and semi-dry electrodes are engineered to overcome these barriers. Recent advancements focus on two key areas:

1. **Materials and Mechanics:** Modern dry sensors utilize novel materials like conductive polymers, flexible metallic coatings, or nanostructured surfaces to establish a stable electrical contact without gel.[31] To navigate through hair, a major obstacle for scalp EEG, electrodes are often designed with pointed, brush-like, or flexible spiky structures that can make direct contact with the scalp without piercing the skin, enhancing both signal quality and patient comfort.[35] Recent research is even exploring micro-scale sensors that are small enough to be inserted between hair follicles, offering a minimally invasive approach with high-fidelity signal capture for up to 12 hours.[36]
2. **Integrated Electronics (Active Electrodes):** Perhaps the most crucial innovation is the integration of a pre-amplifier directly into the electrode housing, creating an "active electrode".[33] The high impedance of the dry skin-electrode contact makes the weak EEG signal highly susceptible to environmental electrical noise. By amplifying the signal at the point of acquisition, active electrodes convert the high-impedance source into a low-impedance output, dramatically improving the SNR and making the signal more robust to artifacts.[33] Advanced systems employ ultra-high input impedance amplifiers and active noise cancellation circuits to achieve signal quality that is comparable to conventional wet electrodes.[37]

These technological enablers are converging in the development of wearable, high-density EEG arrays. These systems are designed as integrated headsets or caps that can be applied rapidly—often in under five minutes—by personnel with minimal training, akin to putting on a hat.[37] This addresses the critical need for rapid deployment in an acute care setting. Products like the Zeto headset, which is the first FDA-approved true dry electrode EEG system, are specifically designed for clinical applications, offering a wireless, adjustable, and easy-to-use solution.[33]

The development of such systems makes the prospect of technician-free application in the ICU increasingly feasible. The combination of rapid setup, improved patient comfort, and the ability to perform long-term monitoring without gel degradation directly addresses the primary logistical barriers to clinical adoption.[32] While challenges related to motion artifacts and ensuring consistent contact pressure remain [32], the clear trajectory is toward robust, user-friendly hardware. This hardware evolution is not merely an incremental improvement; it is a foundational enabler for the entire non-invasive BCI paradigm. By allowing for the collection of long-term, continuous data in a patient's real-world environment, these wearable systems can generate the massive, longitudinal datasets required to train the powerful, data-hungry deep learning models that represent the next frontier in neural decoding. This creates a powerful synergistic feedback loop where hardware advancements directly facilitate progress in software and algorithms.

Section 2: From Raw Signal to Semantic Intent: Advanced Processing and Feature Extraction

Once a neural signal is acquired, the next formidable challenge is to extract the specific patterns corresponding to speech intent while discarding the vast amount of confounding noise. This section explores the "software" side of this problem, examining the known neural correlates of imagined speech, the critical need for sophisticated artifact suppression in the ICU, and the ongoing paradigm shift from manual feature engineering to automated feature discovery with deep learning. The inherent subtlety and variability of imagined speech signatures necessitate a move away from generalized models toward highly personalized systems that can adapt to an individual's unique neural landscape.

2.1. Uncovering the Signatures of Thought: Neural Correlates of Imagined Speech in EEG

Decoding imagined speech requires identifying its unique neural signature. Research suggests that imagined speech production shares significant neural infrastructure with overt, articulated speech, but with crucial differences. It is often conceptualized as an interrupted form of articulation, engaging language-related areas (like Broca's and Wernicke's areas) and the sensorimotor cortex, but stopping short of final motor execution.[6] Functional neuroimaging studies confirm a substantial overlap in activation patterns, although some research suggests that right-hemisphere motor areas may be specifically involved in the active *inhibition* of vocalization during speech imagery.[41]

The neural information related to speech is encoded across various frequency bands in the EEG signal, each associated with different aspects of processing:

- **High-Gamma Band (70–150 Hz):** This band shows strong power modulation during both overt and covert speech production and is considered highly informative, particularly in high-fidelity invasive recordings.[1]
- **Theta Band (4–8 Hz):** Theta oscillations are closely linked to the temporal structure of speech, such as syllable-level processing, phonemic restoration, and have also been associated with the vividness of inner speech.[1]
- **Beta Band (13–30 Hz):** Beta activity is classically associated with motor planning, execution, and feedback, playing a key role in the preparatory stages of speech.[1]
- **Alpha Band (8–12 Hz):** Alpha oscillations are implicated in attentional mechanisms, auditory feedback, and top-down inhibition during language processing.[1]
- **Delta Band (<4 Hz):** The slowest brain waves, delta oscillations, appear to track the prosodic and rhythmic contours of speech, such as intonation and stress patterns.[1]

In addition to frequency-specific power changes, Event-Related Potentials (ERPs)—stereotyped voltage deflections time-locked to a stimulus—also provide clues. While ERPs like the P300 are the basis for many BCI spellers, their direct modulation by the semantic or phonetic content of *imagined* speech is an area of active research. Studies show that the sequence of ERP components (e.g., P100, N200, P300, N400) is similar for both imagined and articulated speech, suggesting a common cognitive processing cascade from perception to semantic integration.[44]

The primary challenge is that these neural signatures—whether in frequency bands or ERPs—exhibit tremendous variability, not only between different individuals but also within the same individual across different sessions or even different trials.[4] This high degree of non-stationarity means that a "one-size-fits-all" decoding model based on a fixed set of features is unlikely to be robust. This inherent variability logically leads to the conclusion that successful speech BCI systems must be highly personalized, capable of learning and adapting to the unique and dynamic neural patterns of each user. This necessity for personalization directly informs the architectural choices and learning paradigms discussed in subsequent sections, particularly the drive toward few-shot learning.

2.2. Signal in the Noise: Artifact Suppression Strategies for the Acute Care Environment

The ICU is an exceptionally "electrically hostile" environment for EEG recording, presenting an extreme version of the signal-to-noise problem.[39] The weak neural signals of interest are inundated by two main categories of artifacts:

1. **Non-Physiological Artifacts:** These originate from the myriad of electronic equipment essential for patient care. Ventilators, infusion pumps, patient warming/cooling blankets, dialysis machines, and even the standard 50/60 Hz powerline interference create a cacophony of electrical noise that can contaminate the EEG recording.[46]
2. **Physiological Artifacts:** These are electrical signals generated by the patient's own body but are not of cortical origin. In the ICU, these are particularly problematic and include electrooculographic (EOG) signals from eye blinks and movements, and electromyogenic (EMG) signals from muscle activity. Patient care activities such as suctioning, repositioning, or performing a sternal rub can generate large, rhythmic artifacts that can easily be mistaken for pathological brain activity like seizures.[39]

Traditional artifact removal methods have significant limitations in this context. Simple digital filters (e.g., band-pass or notch filters) can remove well-defined noise like powerline interference, but they are ineffective against physiological artifacts like EMG, which have a broad frequency spectrum that heavily overlaps with the neural signals of interest.[7] Overzealous filtering risks removing valuable brain data along with the noise.[46]

Blind Source Separation (BSS) techniques, most notably Independent Component Analysis (ICA), have been the gold standard for artifact removal in research settings.[49] ICA decomposes the multi-channel EEG data into a set of statistically independent source signals. An expert can then visually inspect these components, identify those corresponding to artifacts (e.g., eye blinks, muscle tension), and remove them before reconstructing a "clean" EEG signal.[51] While powerful, this approach is ill-suited for a real-time ICU application. It is computationally intensive, often requires manual intervention and expert supervision, and is not designed for online, continuous processing.[53]

To overcome these limitations, the field is rapidly moving toward automated artifact suppression using deep learning. These methods typically employ an autoencoder architecture, where a neural network learns to reconstruct a clean EEG signal from a noisy input. Various architectures are being explored, including Convolutional Neural Networks (CNNs) that are effective at removing specific artifact types like EMG (e.g., NovelCNN), Recurrent Neural Networks (RNNs) that capture temporal features, and Transformers that model both local and non-local dependencies.[49] Recent models like LSTEEG and DeepSeparator are designed as end-to-end frameworks that can automatically detect and correct a variety of artifacts in real-time with minimal delay, a critical requirement for a functional BCI.[53]

However, a truly intelligent BCI for the ICU may require a more nuanced approach than simply rejecting all artifacts. A muscle artifact from a facial grimace is not just noise; it is a potential communication of pain. An ideal system would not just discard this signal but would have a sophisticated, context-aware suppression module. Such a module would function as a classifier in its own right, learning to differentiate between (a) irrelevant environmental noise to be discarded, (b) neural signals related to speech intent to be passed to the decoder, and (c) non-speech physiological signals that are clinically relevant and should be flagged for caregiver attention. This transforms artifact suppression from a simple pre-processing step into an integral, co-adaptive component of the clinical monitoring system itself.

2.3. The End-to-End Paradigm Shift: Feature Engineering vs. Direct Deep Learning

The method by which discriminative information is extracted from the EEG signal has undergone a fundamental paradigm shift, moving away from manual feature engineering toward automated, end-to-end deep learning.

The classical approach involves a multi-stage pipeline built on domain expertise.[5] First, the raw EEG signal is heavily pre-processed and filtered. Then, a human expert selects and engineers a set of features believed to be relevant for the task. These features are typically extracted from different domains: the time domain (e.g., signal amplitude at specific latencies), the frequency domain (e.g., power within specific bands like alpha or beta), or the time-frequency domain (e.g., using wavelet transforms to see how frequency content changes over time).[5] This curated set of features is then fed into a standard machine learning classifier, such as a Support Vector Machine (SVM) or a Random Forest, to make a prediction.[4] The main advantage of this approach is its interpretability and the ability to incorporate decades of neuroscientific knowledge into the model. However, it is a laborious, time-consuming process that is highly dependent on the skill of the expert. If the chosen features are suboptimal, or if crucial information is discarded during pre-processing, the performance of the entire system is capped, a phenomenon sometimes referred to as "garbage in, garbage out".[58]

In contrast, the end-to-end deep learning paradigm seeks to automate this process.[56] Deep neural networks, with their multiple layers of processing, are capable of learning a hierarchy of features directly from raw or minimally processed data.[60] The initial layers of the network might learn to identify simple patterns, while deeper layers combine these to form more abstract and complex representations relevant for the final classification or synthesis task. This approach significantly simplifies the development pipeline and holds the potential to discover novel, highly discriminative features that a human expert might not have considered.[58]

This power comes at a cost. End-to-end models are notoriously data-hungry; they require massive datasets to learn these features effectively and avoid overfitting.[56] They are also computationally expensive to train. A more significant barrier for clinical adoption is their "black box" nature. It can be extremely difficult to understand

why a deep learning model made a particular decision, which poses a challenge for validation and trust in a high-stakes medical environment.[56] Despite these challenges, the superior performance and scalability of end-to-end models have made them the dominant approach in modern BCI research, with ongoing efforts focused on improving their data efficiency and interpretability.

Section 3: Architectures of a Synthetic Voice: The Evolution of Neural Decoding Models

At the heart of any speech BCI is the decoder: a machine learning model tasked with the formidable challenge of translating complex, noisy patterns of brain activity into a representation of language. This section traces the rapid evolution of these decoding architectures, from early linear models to the sophisticated deep neural networks in use today. It further analyzes the critical design choice of what to decode—discrete words, phonetic building blocks, or continuous acoustic waveforms—and examines the pressing need to overcome the "calibration hurdle" to make these systems clinically practical.

3.1. From Linear Models to Transformers: A Trajectory of Time-Series Decoders

The sophistication of neural decoders has grown in lockstep with advances in the broader field of machine learning for time-series analysis. Early BCI systems often relied on linear models or classical machine learning algorithms like Support Vector Machines (SVMs) applied to hand-engineered features.[1] While useful for simple classification, these models lack the capacity to capture the intricate, non-linear, and dynamic nature of neural data associated with a complex process like speech.

The first major leap forward came with the application of Recurrent Neural Networks (RNNs), especially variants like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs).[4] RNNs are explicitly designed to process sequential data. Their architecture includes a "hidden state" or memory that allows them to retain information from previous time steps, making them naturally suited for decoding EEG signals, which unfold continuously over time.[64] Indeed, some of the most successful invasive speech BCIs have used RNNs to achieve high-accuracy, real-time text decoding.[14] However, RNNs process data sequentially, which can make them computationally slow and prone to issues like the "vanishing gradient problem," where their ability to learn long-range dependencies diminishes.[64]

Temporal Convolutional Networks (TCNs) emerged as an alternative, using techniques from convolutional networks (CNNs) to process time-series data.[66] By using dilated convolutions, TCNs can achieve a very large receptive field, allowing them to see long stretches of historical data. They are also fully parallelizable, making them faster to train than RNNs.

The most recent and impactful revolution has been the adaptation of the Transformer architecture for neural decoding.[67] Originally designed for natural language translation, the Transformer's power lies in its

self-attention mechanism.[69] Unlike an RNN, which processes a sequence one step at a time, the attention mechanism allows the model to weigh the importance of every time point in the input sequence relative to every other time point, all at once. This ability to capture complex, non-local, and long-range dependencies is exceptionally well-suited to modeling brain activity, which is not a simple linear process but the result of a highly parallel, interconnected network.[70] The success of Transformers suggests that the neural code for speech is itself a high-dimensional representation where context from distant points in time is critically important. To leverage the strengths of different approaches, state-of-the-art research is now focused on hybrid models, which might use convolutional layers to extract local spatiotemporal features and then feed these into a Transformer to model the global relationships between them.[5]

Table 2: Evolution of Speech Decoding Architectures

Architecture	Key Characteristics	Strengths for Neural Time-Series	Weaknesses/Challenges	Representative Studies
RNN (LSTM/GRU)	Sequential processing, internal memory state (gates) to regulate information flow.	Captures temporal dependencies in sequential data like EEG. Effective for modeling signals that unfold over time.	Vanishing/exploding gradients over long sequences; sequential nature limits parallelization.	[4]
TCN	Uses causal, dilated convolutions to process time-series data.	Large receptive field, parallelizable, more stable training than RNNs.	May be less effective at capturing very long-range or complex non-local dependencies compared to Transformers.	[66]
Transformer	Parallel processing via self-attention mechanism; positional encodings maintain sequence order.	Excellently captures long-range and global dependencies; highly parallelizable, leading to faster training on large datasets.	Data-hungry, computationally expensive, may overlook fine-grained local features without a convolutional component.	[67]

Hybrid (CNN+Transformer)	Combines convolutional layers for local/spatial feature extraction with a Transformer encoder for global dependency modeling.	Leverages the best of both worlds: CNNs find local patterns, and Transformers model the global context between them.	Increased model complexity.	5

3.2. The Target Representation Dilemma: Classifying Words, Phonemes, or Synthesizing Acoustics

A crucial design choice for any speech BCI is the nature of its output. There are three primary strategies, each with a distinct set of advantages and limitations that reflect a trade-off between simplicity, scalability, and naturalness.

1. **Direct Classification:** This is the most straightforward approach, treating the BCI as a classifier that decodes a complete word or phrase from a small, predefined list.[5] For example, the system might be trained to recognize the neural patterns for "Yes," "No," "Pain," and "Water." This method is computationally simpler and has been the foundation of most non-invasive BCI research to date. Its fatal flaw, however, is its lack of scalability. A system limited to a vocabulary of 5, 10, or even 50 words cannot support the open-ended, spontaneous communication needed for true interaction.[1]
2. **Phonetic/Articulatory Decoding:** A far more powerful and scalable approach is to decode the fundamental building blocks of speech rather than whole words. This involves training the decoder to identify the neural correlates of either phonemes (the basic sounds of a language, like /b/, /a/, /t/) or the underlying motor commands sent to the articulators (lips, tongue, jaw, larynx) to produce those sounds.[72] With a set of just ~40 phonemes, it is possible to construct nearly any word in the English language. This is the strategy employed by the most successful high-performance invasive BCIs.[11] In these systems, the neural decoder outputs a stream of probabilities for each phoneme over time. This stream is then fed into a statistical language model (similar to the one in a smartphone's autocorrect) which determines the most likely sequence of words,

effectively correcting errors and imposing grammatical structure.[11] This approach provides a pathway to a massive, open vocabulary.

3. **Acoustic Synthesis:** The most ambitious and potentially most natural approach is to bypass discrete units like phonemes or words altogether and directly synthesize a continuous acoustic representation of speech, such as a spectrogram or an audio waveform, from brain activity.[74] The advantage of this method is its potential to reconstruct not just the linguistic content (the words) but also the paralinguistic features of speech—the prosody, intonation, emotion, and rhythm that carry a huge portion of communicative meaning.[61] By using a model of the patient's pre-injury voice, it's even possible to synthesize speech that sounds like the user, a factor that has been shown to increase the sense of embodiment and agency.[63] However, this is an exceptionally difficult task, especially with noisy non-invasive signals. Recent pioneering work with end-to-end frameworks like FESDE aims to map EEG signals directly to audio waveforms, which in theory preserves more of the rich acoustic information that is lost when converting to text as an intermediate step.[59]

The choice between these targets is more than a technical decision; it reflects a philosophical difference in the goal of the BCI. Direct classification creates an *alternative* communication channel, functionally similar to an advanced assistive switch.[75] Phonetic and acoustic synthesis, on the other hand, aim to tap into and *restore* the brain's natural speech production pathway. The latter holds the promise not just of communication, but of restoring a fundamental component of a person's identity and enabling a more natural and embodied form of interaction.

3.3. The Calibration Hurdle: Toward "Zero-Shot" and "Few-Shot" BCI Deployment

A major practical barrier preventing the widespread clinical adoption of BCI technology is the immense burden of calibration.[30] Most high-performance decoders are highly personalized and require a substantial amount of patient-specific training data to function accurately. This calibration process can involve many hours or even days of a patient repeatedly attempting to speak words or phrases while their neural activity is recorded.[17] This is impractical and often impossible for a patient in an acute care setting, who may be fatigued, in pain, or have a fluctuating level of consciousness. For a speech BCI to be clinically viable in the ICU, it must be deployable and functional within minutes, not days.

This clinical imperative has fueled intense research into **zero-shot** and **few-shot learning** paradigms, which aim to create models that can generalize to new users or new tasks with little to no specific training data.[77]

- **Zero-Shot Learning:** In this approach, the model is not trained to map neural signals to specific, pre-defined words. Instead, it learns to map brain activity to a continuous, shared "embedding space" that describes the properties of the stimulus. For example, it might learn the relationship between brain activity and a set of semantic features that define a word. The system could then, in theory, decode a word it has never seen before by identifying its corresponding location in this semantic space.[80]
- **Few-Shot Learning:** This is arguably the more practical approach for near-term clinical use. Here, a large, general model is pre-trained on a massive dataset, which could include data from many other subjects or even other BCI tasks. This pre-trained model is then rapidly adapted or fine-tuned to a new patient using only a very small amount of their specific data—a "few shots".[82]

The key enabling technologies for this are **transfer learning** and **domain adaptation**. These techniques allow knowledge gained in one domain (the "source," e.g., a large public dataset) to be transferred to another (the "target," e.g., the new ICU patient).[40] For instance, a model could be pre-trained on a patient's own articulated speech signals (if collected before they lost the ability to speak) and then adapted to decode their imagined speech signals.[45] Another powerful technique is Model-Agnostic Meta-Learning (MAML), which specifically trains a model's initial parameters to be in a state where they can be rapidly adapted to a new task with only a few examples and minimal computational updates.[82] The development of standardized benchmark datasets, such as the recently introduced FALCON suite, is critical for systematically evaluating and comparing these few-shot and zero-shot approaches to ensure they are robust for real-world use.[76]

Section 4: From Lab Bench to ICU Bedside: Clinical Translation and Human-Centered Design

The successful translation of a speech BCI from a controlled laboratory environment to the chaotic reality of an ICU requires overcoming a gauntlet of practical challenges that extend far beyond the core decoding algorithm. This section examines the critical human and environmental factors that govern clinical viability. It explores the unique physiological and cognitive barriers presented by critically ill patients, outlines principles for designing usable interfaces for both patients and caregivers, and establishes a framework for the robust performance benchmarks that must be met to define clinical success.

4.1. The Translational Gauntlet: Overcoming Patient and Environmental Barriers in Critical Care

Deploying a BCI in the ICU means shifting from healthy, motivated volunteers to a patient population characterized by profound physiological and cognitive instability. These patient-specific factors represent a primary translational hurdle:

- **Fatigue and Cognitive Load:** Using a BCI is an inherently demanding mental task.[85] ICU patients suffer from extreme fatigue due to their illness, sleep disruption, and the stressful environment. Their cognitive resources are already heavily taxed, and they may have fluctuating levels of consciousness.[87] This high cognitive load can severely impair their ability to generate the consistent, focused neural signals required for BCI control, leading to performance degradation over time.[85] Studies have shown a significant relationship between self-reported fatigue and frustration and BCI performance.[85] Visual fatigue is a particularly significant factor for systems that rely on visual stimuli, like SSVEP-based BCIs.[87] Managing this cognitive load is therefore a central challenge for system design.[89]
- **Concurrent Medical Interventions:** The brain activity of an ICU patient is not a stable baseline. It is constantly modulated by the effects of sedatives, analgesics, and other medications, as well as by the underlying pathophysiology of their condition.[90] These factors can introduce significant non-stationarity into the EEG signals, making it difficult for a decoder trained at one point in time to remain accurate.
- **Physical and Sensory Impairments:** Beyond the inability to speak, patients may have other impairments that limit BCI options. For example, eyelid apraxia (the inability to voluntarily open the eyes) would render any BCI requiring visual attention unusable.[90] Pain, skin breakdown, and peripheral neuropathy can make tactile-based systems challenging.[90]

The ICU environment itself poses additional barriers. It is a setting of constant auditory and physical distractions, with alarms, staff conversations, and patient care activities creating an environment that is antithetical to the quiet focus often required for BCI use.[90] As detailed in Section 2.2, it is also an electrically hostile environment, demanding extremely robust hardware and artifact suppression algorithms.[46]

These challenges necessitate a move from the synchronous, cue-based paradigms common in laboratory research to more flexible, self-paced systems. A cue-based system, which only allows the user to communicate during predefined time windows, is too rigid for a patient with a fluctuating mental state.[90] A self-paced (or asynchronous) system, which is always "on" and allows the user to initiate communication at will, is far more practical and empowering.[90] However, self-paced systems introduce the difficult challenge of the "intent detection" problem: the system must reliably distinguish between periods when the user intends to communicate and periods of rest, with a very low rate of false positives to avoid generating unwanted output.[90]

A critical and often overlooked human factor is the concept of "BCI illiteracy".[92] Not every individual can learn to proficiently control a BCI. Studies on motor imagery BCIs have shown that a significant portion of users (up to 50% in some cases) are unable to achieve control above chance level, and similar variability is seen in speech imagery BCIs.[43] This phenomenon is not fully understood but likely relates to individual differences in neuroanatomy, cognitive strategies, and attentional abilities. This has profound implications for clinical translation. It means a BCI cannot be prescribed with a predictable outcome like a pharmaceutical drug; its success depends on the patient's ability to learn a complex new skill. This reality complicates clinical trial design, resource allocation (as a significant percentage of patients may not benefit), and, critically, the process of managing patient and family expectations.

4.2. The Human-in-the-Loop: Principles of Usability and Interface Design for Patients and Caregivers

For a BCI to be adopted in a clinical setting, it must be not only accurate but also usable. This has prompted a shift in the field away from a narrow focus on technical metrics like accuracy and information transfer rate (ITR) toward a more holistic, user-centered design (UCD) approach that prioritizes effectiveness, efficiency, and user satisfaction.[93]

The Patient Interface: The user interface (UI) for the patient must be designed with the explicit goal of minimizing extraneous cognitive load.

- **Feedback is Paramount:** The single most important principle is the provision of clear, intuitive, and instantaneous feedback. The user cannot learn to modulate their brain signals without knowing the effect of their mental effort. This closed-loop neurofeedback is essential for skill acquisition.[95] Research has shown that multimodal feedback—for example, combining a visual display with auditory confirmation—leads to significantly greater performance in learning to control a BCI than either visual or auditory feedback alone.[96] A simple but effective UI element could be a visual gauge on a screen that fills up in proportion to the clarity of the detected speech imagery, giving the user real-time insight into their performance.[97]
- **Mental Task Selection:** The choice of the mental task used to generate the control signal is also a key UI/UX consideration. While motor imagery (e.g., imagining hand movement) is a common paradigm, it may not be optimal for all users. Exploring alternative paradigms like imagining spatial navigation (e.g., moving through a familiar room) or imagining a piece of music may provide more effective or less fatiguing control signals for certain individuals.[98]

The Caregiver Interface: The interface for the clinical staff is just as critical as the patient's. A BCI that adds to the already immense workload of an ICU nurse will not be used.

- **Clarity and Confidence:** The primary output—the decoded speech—must be presented in a clear, unambiguous way. Crucially, the interface should not just display the decoder's single best guess; it must also convey the system's confidence in that decoding. This allows the caregiver to triage the information, paying immediate attention to high-confidence messages while seeking clarification for low-confidence ones.
- **Workflow Integration and Alerting:** The system must seamlessly integrate into the existing clinical workflow. It should provide a clear indication of its operational status (e.g., online, calibrating, offline). As previously discussed, an intelligent system could also serve as an alert mechanism, flagging clinically relevant physiological artifacts (e.g., signs of pain or distress) that might otherwise go unnoticed.

The evaluation of these systems must extend beyond technical benchmarks. Standardized subjective measures, such as the NASA Task Load Index (NASA-TLX), should be used to quantify the perceived mental, physical, and temporal demand, effort, and frustration associated with using the BCI.[94] This provides a crucial measure of efficiency from the user's perspective, which is often more important for acceptance than raw speed.

Finally, there is a fundamental tension in design goals between a BCI intended for an *acute* setting like the ICU and one designed for *chronic* use at home. An ICU system must prioritize rapid, technician-free deployment, extreme robustness to noise, and the decoding of a limited but critical vocabulary related to immediate care needs (e.g., "pain," "suction," "family"). The primary consumer of the output is the clinical team. In contrast, a chronic home-use system might prioritize long-term comfort, cosmetic appeal, and open-vocabulary communication for social connection and environmental control (e.g., controlling a smart home).[99] This suggests that a "one-size-fits-all" speech BCI is an unrealistic goal; instead, technology must be tailored to the specific context of use.

4.3. Defining Success: Establishing Robust Performance Benchmarks for Clinical Utility

To assess clinical viability, the field must move beyond simple classification accuracy and adopt a set of robust, standardized performance benchmarks that reflect the real-world utility of a communication device. The key metrics are speed, accuracy on conversational-level tasks, and, for synthesis systems, intelligibility.

- **Words Per Minute (WPM):** This metric quantifies the speed of communication, a critical factor for usability. Slow, laborious communication can lead to frustration and abandonment. For context, natural human conversation occurs at approximately 160 WPM.[15] The current record for an invasive speech BCI is an impressive 78 WPM, which is more than four times faster than previous records and begins to approach a functional conversational speed.[12] This is the benchmark that non-invasive systems must ultimately strive for. For an ICU patient, a clinically useful speed would be one that is demonstrably faster and less effortful than low-tech alternatives like partner-assisted letter scanning.
- **Word Error Rate (WER):** WER is the gold standard for evaluating the accuracy of automated speech recognition (ASR) and is the most meaningful accuracy metric for speech BCIs.[100] It is calculated as the sum of word substitutions, deletions, and insertions, divided by the total number of words in the ground-truth reference transcript:
  $WER=(S+D+I)/N$. A lower WER indicates higher accuracy. State-of-the-art invasive

systems have achieved WERs as low as 9.1% on a constrained 50-word vocabulary and 23.8% on a very large 125,000-word vocabulary.[15] The recent Brain-to-Text Benchmark '24 challenge, which provides a standardized platform for comparing algorithms, saw the winning entry achieve a 5.8% WER on a held-out dataset, a substantial improvement over the 9.7% baseline, demonstrating rapid progress in decoder optimization.[102]

- **Short-Time Objective Intelligibility (STOI):** For BCIs that directly synthesize an acoustic waveform, WER is not applicable. Instead, objective intelligibility metrics are used. STOI is one such metric that measures the correlation between the temporal envelopes of the original and reconstructed speech signals.[65] It produces a score between -1 and 1, where higher values indicate better intelligibility and have been shown to correlate well with human listener perception.

Establishing a clear threshold for "clinical utility" is essential. While no universal standard exists, a WER above 10-20% can be highly problematic for effective communication.[101] For a non-invasive system to be considered viable in an ICU, it would need to achieve a combination of acceptable speed and accuracy on a vocabulary that is functionally useful for that environment.

Table 3: Clinical Utility Benchmarks for Speech BCIs

Metric	Definition	Current SOTA (Invasive)	Current SOTA (Non-Invasive)	Proposed Threshold for ICU Clinical Utility
Words Per Minute (WPM)	Speed of decoded communication.	78 WPM [13]	Primarily classification-based; not a standard metric yet.	>15 WPM (Faster than partner-assisted scanning)
Word Error Rate (WER)	% of incorrectly decoded words.	5.8% (125k vocab, benchmark) [102]; 9.1% (50 vocab) [15]	High; not yet benchmarked for open vocab.	<30% on a core vocabulary of >100 words.

Short-Time Objective Intelligibility (STOI)	Objective measure of synthesized speech intelligibility (-1 to 1).	Not a primary metric for text-based decoders.	Still in early research (e.g., FESDE).[59]	>0.75 (for acoustic synthesis systems)
Calibration Time	Time required for patient-specific training before use.	Days to weeks of sessions.	Hours to days.	<15 minutes for initial functional use.

Section 5: The Neuroethical Imperative: Privacy, Consent, and Responsibility in Speech BCI

The development of technologies capable of interpreting neural signals related to thought and language carries profound ethical responsibilities. As speech BCIs transition from theoretical concepts to clinical tools, their successful and ethical implementation hinges on proactively addressing complex issues of privacy, autonomy, consent, and the consequences of error. The ethical considerations are evolving from a traditional "protection-from-harm" model, focused on physical risks and data security, to a more nuanced "promotion-of-well-being" framework that must grapple with complex concepts of identity, agency, and the very nature of the self.

5.1. The Sanctity of Thought: Neuro-privacy and the Challenge of Informed Consent

BCI technology introduces privacy risks that are fundamentally different from those associated with other digital technologies. The concern is not merely for data privacy, but for **mental privacy**: the right to control access to one's own thoughts, emotions, and inner monologue.[103] Neural data, unlike data we voluntarily input into a system, can potentially reveal unfiltered, subconscious, or unintentional cognitive states.[104] This raises the specter of "thought surveillance" or "cognitive hacking," where neural data could be exploited for commercial targeting, psychological warfare, or blackmail.[106] While current technology is far from this scenario, the ethical framework must anticipate it.

This leads to one of the most acute ethical challenges: obtaining meaningful **informed consent** from the very population the technology aims to serve—non-verbal individuals with severe motor impairments.[108] The standard process of consent assumes a person can communicate their understanding and decision. For a locked-in patient, this is immensely difficult. Key challenges and best practices include:

- **Assessing Decisional Capacity:** It is extraordinarily difficult to determine if a patient who cannot speak or move fully comprehends the complex risks, benefits, and alternatives of participating in BCI research.[110] The very tool needed to express consent is what the research is trying to build.
- **Role of Surrogates and Assent:** While obtaining consent from a legally authorized representative is standard procedure, every effort must be made to involve the patient directly. This includes seeking the patient's *assent* (agreement) through any reliable communication channel they possess, such as eye blinks or residual movements. Crucially, a patient's *dissent*, however weakly signaled, must be respected.[110]
- **Dynamic and Continuous Consent:** Consent cannot be a one-time event at the start of a study. Given the fluctuating condition of ICU patients and the evolving nature of the research, consent must be an ongoing process. Researchers should repeatedly check in with the patient to ensure their continued willingness to participate.[108]
- **Clarity and Accessibility:** Information must be communicated in the clearest, most accessible manner possible, avoiding technical jargon and using visual aids or simplified summaries to explain the technology, procedures, potential risks (e.g., skin irritation from electrodes, frustration), and realistic benefits.[108]

The legal landscape for these "neuro-rights" is only just beginning to form. Chile has taken the pioneering step of enshrining neuro-rights in its constitution, and some jurisdictions in the United States are starting to explicitly include "neurological data" under existing privacy laws, but comprehensive, specific regulation remains absent.[104]

5.2. The Burden of Hope: Managing Patient and Familial Expectations

For patients and families facing devastating neurological conditions, the prospect of a speech BCI can generate immense hope. This places a heavy ethical burden on researchers and clinicians to manage expectations responsibly and avoid fostering unrealistic beliefs that can lead to disappointment and distress.[105] Media portrayals often present a sensationalized view of BCI capabilities, which can be difficult to counteract.[110]

Effective management of expectations requires a shift in focus from the technology itself to the patient's overall well-being and social context. The "disability paradox" observes that an individual's quality of life is often more strongly correlated with the quality of their social relationships and support systems than with their level of physical impairment.[113] Therefore, a BCI should be framed not as a magic cure, but as one tool among many to help maintain and enhance those crucial connections.

Key strategies for responsible communication include:

- **Collaborative and Realistic Goal Setting:** The decision to use a BCI should be made by a collaborative team that includes the patient, family, and clinical staff.[114] The goals for the technology should be concrete and functional (e.g., "to be able to answer yes/no questions reliably," "to communicate pain level") rather than abstract promises of "restoring natural speech".[114]
- **Honest and Transparent Communication:** Clinicians must be forthright about the limitations of the technology, including the potential for it not to work for a particular patient (BCI illiteracy), the high cognitive effort required, and the current performance benchmarks.[112]
- **Comprehensive Family Support:** The introduction of a complex new technology affects the entire family unit. The care plan must include education, training, and emotional support for caregivers, who are often under immense stress and may need help managing their own expectations and avoiding burnout.[112]

5.3. When the Signal is Wrong: Mitigating the Consequences of Misinterpretation and Error

All BCI systems are imperfect and will make decoding errors. In the context of a speech BCI for a critically ill patient, an error is not a trivial matter. An incorrect word could lead to a catastrophic miscommunication regarding a medical decision (e.g., consent for a procedure, end-of-life wishes), a failure to report pain, or damage to personal relationships.[103]

These errors raise difficult questions of **autonomy and responsibility**. If a BCI, aided by a predictive language model, produces an utterance that the user did not intend, who is the author of that statement? Who is held responsible for its consequences? This ambiguity can erode the user's sense of agency and "first-person authority," the feeling that they are in control of their own expression.[103]

Mitigating these risks requires that safety and uncertainty management be designed into the system from the ground up:

- **Quantifying and Flagging Uncertainty:** A safe BCI should not simply output its single best guess. The underlying algorithms must produce a confidence score for each decoding. The user interface should then clearly flag low-confidence outputs to both the user and the caregiver, signaling that the message may be incorrect and requires verification.
- **Implementing User Veto Control:** The user must have ultimate authority over the output. The system must incorporate a simple, low-effort, and reliable mechanism for the user to cancel or veto an incorrect decoding before it is fully communicated to others. This is a non-negotiable feature for preserving user autonomy.[103]
- **Context-Aware Safeguards:** For communication with high-stakes consequences, the system could be designed with additional safeguards. For example, when asking about a critical medical decision, the system might require a higher confidence threshold to be met or engage in a multi-step verification process to confirm the user's intent.

Finally, the advancement of BCI technology raises a broader societal concern. The most acute clinical need for this technology is within a relatively small patient population with severe paralysis.[11] However, a much larger and more lucrative potential market exists in cognitive enhancement for healthy individuals, a long-term goal explicitly stated by some commercial BCI developers.[13] This creates a risk that research and investment could pivot away from the difficult, less profitable therapeutic applications toward consumer enhancement products. This potential "neuro-enhancement divide" poses a significant challenge to social justice and the equitable distribution of transformative medical technologies, and it requires careful consideration by policymakers, funders, and the research community to ensure that the needs of the most vulnerable patients remain a priority.[13]

Conclusion

The prospect of a non-invasive Brain-Computer Interface capable of restoring real-time, intelligible speech to non-verbal patients in an acute care setting remains on the horizon, not yet at the bedside. The analysis reveals a field defined by a stark performance chasm between high-risk, high-reward invasive technologies and safer, but currently underperforming, non-invasive approaches. For a technology based on scalp EEG to become clinically viable in the ICU, it must overcome fundamental challenges rooted in physics, engineering, and human biology.

Despite these formidable obstacles, the trajectory is one of accelerating progress, fueled by a powerful convergence of innovation across three critical fronts:

1. **Hardware Maturation:** The development of wearable, dry-electrode EEG systems with integrated active electronics is a cornerstone advancement. These systems are making non-invasive BCI more practical, comfortable, and suitable for the long-term, continuous data acquisition required in a clinical environment.
2. **Algorithmic Sophistication:** The paradigm shift to end-to-end deep learning, particularly with complex architectures like Transformers, provides the necessary horsepower to begin untangling the subtle neural signatures of imagined speech from noisy data.
3. **Data Efficiency:** The emergence of few-shot and zero-shot learning techniques offers a credible path to overcoming the crippling burden of patient-specific calibration, a prerequisite for any system intended for rapid deployment in acute care.

The path forward, however, is not purely technical. The greatest remaining challenges are translational and ethical. The unique complexities of the ICU patient—physiological instability, cognitive load, and fluctuating consciousness—demand a new level of system robustness and adaptability. Success will be contingent on adopting a deeply human-centered design philosophy that prioritizes usability for both patients and caregivers, and on establishing clear, clinically meaningful benchmarks for performance that go beyond laboratory metrics.

Ultimately, the development of speech BCIs must proceed with a profound sense of ethical responsibility. The journey toward decoding thought necessitates the creation of robust frameworks for protecting mental privacy, ensuring meaningful informed consent can be obtained from those who cannot speak, managing the heavy burden of hope placed on patients and families, and designing systems that are fundamentally safe and accountable. While the technical challenges are immense, they are matched by the potential to restore a fundamental aspect of human dignity and connection. Achieving this goal will require a concerted, interdisciplinary effort that inextricably links scientific discovery with clinical wisdom and ethical foresight.

Works cited

1. A State-of-the-Art Review of EEG-Based Imagined Speech Decoding - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC9086783/
2. EEG-Based BCIs on Motor Imagery Paradigm Using Wearable Technologies: A Systematic Review - PubMed Central, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC10007053/
3. Decoding Covert Speech From EEG-A Comprehensive Review - PMC - PubMed Central, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC8116487/
4. Speech Decoding from EEG Signals - The Science and Information (SAI) Organization, accessed June 28, 2025, https://thesai.org/Downloads/Volume16No4/Paper_85-Speech_Decoding_from_EEG_Signals.pdf
5. Decoding Imagined Speech from EEG Data: A Hybrid Deep Learning Approach to Capturing Spatial and Temporal Features - MDPI, accessed June 28, 2025, https://www.mdpi.com/2075-1729/14/11/1501
6. Toward an Imagined Speech-Based Brain Computer Interface Using EEG Signals - White Rose eTheses Online, accessed June 28, 2025, https://etheses.whiterose.ac.uk/id/eprint/24588/1/thesis.pdf
7. Using articulated speech EEG signals for imagined speech decoding - ISCA Archive, accessed June 28, 2025, https://www.isca-archive.org/interspeech_2024/bras24_interspeech.pdf
8. Modulating Brain Activity with Invasive Brain–Computer Interface: A Narrative Review - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC9856340/
9. Decoding Movement From Electrocorticographic Activity: A Review - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/journals/neuroinformatics/articles/10.3389/fninf.2019.00074/full
10. High-density electrocorticography for speech decoding and ..., accessed June 28, 2025, https://www.news-medical.net/news/20231108/High-density-electrocorticography-for-speech-decoding-and-improved-neural-speech-prostheses.aspx
11. Brain-to-Text Benchmark '24 - Overview - EvalAI, accessed June 28, 2025, https://eval.ai/web/challenges/challenge-page/2099/overview
12. 2nd Place Winner - BCI Award 2022 - A high performance intracortical speech

BCI, accessed June 28, 2025, https://www.youtube.com/watch?v=eEFYb543aLc
13. Ethical considerations for the use of brain–computer interfaces for cognitive enhancement, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC11542783/
14. Brain2Text Dataset - Papers With Code, accessed June 28, 2025, https://paperswithcode.com/dataset/brain2text
15. A high-performance speech neuroprosthesis | bioRxiv, accessed June 28, 2025, https://www.biorxiv.org/content/10.1101/2023.01.21.524489v2.full-text
16. A high-performance speech neuroprosthesis - PMC - PubMed Central, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC10468393/
17. Scaling Law in Neural Data: Non-Invasive Speech Decoding with 175 Hours of EEG Data, accessed June 28, 2025, https://arxiv.org/html/2407.07595v1
18. Decoding Imagined Speech from EEG Data: A Hybrid Deep Learning Approach to Capturing Spatial and Temporal Features - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC11595501/
19. Investigating the Neural Correlates of Imagined Speech: an EEG-based Connectivity Analysis | Request PDF - ResearchGate, accessed June 28, 2025, https://www.researchgate.net/publication/358215835_Investigating_the_Neural_Correlates_of_Imagined_Speech_an_EEG-based_Connectivity_Analysis
20. Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope - PubMed, accessed June 28, 2025, https://pubmed.ncbi.nlm.nih.gov/30205208/
21. Differences between EEG, NIRS, fMRI and MEG - Brain Latam, accessed June 28, 2025, https://brainlatam.com/blog/differences-between-eeg-nirs-fmri-and-meg-924
22. Comparing the potential of MEG and EEG to uncover brain tracking of speech temporal envelope - ResearchGate, accessed June 28, 2025, https://www.researchgate.net/publication/327538186_Comparing_the_potential_of_MEG_and_EEG_to_uncover_brain_tracking_of_speech_temporal_envelope
23. fNIRS vs EEG: Comparing Non-Invasive Brain Imaging Techniques ..., accessed June 28, 2025, https://imotions.com/blog/learning/research-fundamentals/fnirs-vs-eeg/
24. fNIRS-based brain-computer interfaces: a review - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC4309034/
25. Speech Recognition via fNIRS Based Brain Signals - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2018.00695/full
26. Strategic Integration: A Cross-Disciplinary Review of the fNIRS-EEG Dual-Modality Imaging System for Delivering Multimodal Neuroimaging to Applications - MDPI, accessed June 28, 2025, https://www.mdpi.com/2076-3425/14/10/1022
27. g.tec News #1 - Introducing our new products and events in 2025, accessed June 28, 2025, https://www.gtec.at/2025/01/14/introducing-new-products-and-events-in-2025/

28. Towards A Non-Invasive Method to Unlock Brain's Understanding of Speech - USC Viterbi, accessed June 28, 2025, https://viterbischool.usc.edu/news/2025/04/towards-a-non-invasive-method-to-unlock-brains-understanding-of-speech/
29. Invasive Brain–Computer Interface for Communication: A Scoping Review - MDPI, accessed June 28, 2025, https://www.mdpi.com/2076-3425/15/4/336
30. Current Challenges Facing the Translation of Brain Computer ..., accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC4702293/
31. Advancements in dry and semi-dry EEG electrodes ... - AIP Publishing, accessed June 28, 2025, https://pubs.aip.org/aip/adv/article-pdf/doi/10.1063/5.0228644/20501473/040703_1_5.0228644.pdf
32. Dry EEG Electrodes - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC4168519/
33. Comparison of Dry Electrode EEG System with Conventional EEG ..., accessed June 28, 2025, https://zeto-inc.com/blog/comparison-of-dry-electrode-eeg-system-with-conventional-eeg-system/
34. Revolutionizing BCIs: The Future of Dry EEG Sensors - Number Analytics, accessed June 28, 2025, https://www.numberanalytics.com/blog/future-dry-eeg-sensors-bci
35. Advancements in dry and semi-dry EEG electrodes: Design, interface characteristics, and performance evaluation - AIP Publishing, accessed June 28, 2025, https://pubs.aip.org/adv/article/15/4/040703/3345166/Advancements-in-dry-and-semi-dry-EEG-electrodes
36. New Wearable Brain-Computer Interface - Georgia Tech Research, accessed June 28, 2025, https://research.gatech.edu/new-wearable-brain-computer-interface
37. Dry Electrode - Wearable Sensing | Dry EEG, accessed June 28, 2025, https://wearablesensing.com/dry-electrode/
38. Revolutionizing medical care: Wearable EEG-Based BCIs' new ..., accessed June 28, 2025, https://www.eurekalert.org/news-releases/1030488
39. EEG Artifacts in the Intensive Care Unit Setting - ResearchGate, accessed June 28, 2025, https://www.researchgate.net/publication/42973598_EEG_Artifacts_in_the_Intensive_Care_Unit_Setting
40. Towards Voice Reconstruction from EEG during Imagined Speech, accessed June 28, 2025, https://ojs.aaai.org/index.php/AAAI/article/view/25745/25517
41. The neural correlates of the different speech production ... - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/10.3389/conf.neuro.09.2009.01.265/event_abstract
42. Gamma and Theta/Alpha-Band Oscillations in the Electroencephalogram Distinguish the Content of Inner Speech | eNeuro, accessed June 28, 2025, https://www.eneuro.org/content/12/2/ENEURO.0297-24.2025

43. Learning to operate an imagined speech Brain-Computer Interface involves the spatial and frequency tuning of neural activity - PubMed Central, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC11842755/
44. Delft University of Technology Using articulated speech EEG signals for imagined speech decoding - TU Delft Repository, accessed June 28, 2025, https://repository.tudelft.nl/file/File_3f939e38-c08e-471e-8aa4-72ebe947b77d
45. From pronounced to imagined: improving speech decoding with multi-condition EEG data - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/journals/neuroinformatics/articles/10.3389/fninf.2025.1583428/full
46. Electroencephalography and Artifact in the Intensive Care Unit | Neupsy Key, accessed June 28, 2025, https://neupsykey.com/electroencephalography-and-artifact-in-the-intensive-care-unit/
47. A novel EEG artifact in the intensive care unit - PubMed, accessed June 28, 2025, https://pubmed.ncbi.nlm.nih.gov/12477993/
48. accessed January 1, 1970, https.neupsykey.com/electroencephalography-and-artifact-in-the-intensive-care-unit/
49. A novel EEG artifact removal algorithm based on an advanced ..., accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC12134218/
50. Assessment of Artefact Suppression by ICA and Spatial Filtering on Reduced Sets of EEG Signals, accessed June 28, 2025, http://139.91.210.27/CBML/PROCEEDINGS/2011_EMBC/PDFs/Papers/11611592.pdf
51. Ongoing EEG artifact correction using blind source separation - arXiv, accessed June 28, 2025, https://arxiv.org/pdf/2306.16910
52. Independent Components Analysis "Artifact Correction" Distorts EEG Phase in Artifact Free Segments - Sci Forschen, accessed June 28, 2025, https://sciforschenonline.org/journals/neurology/article-data/JNNB172/JNNB172.pdf
53. EEG Artifact Detection and Correction with Deep Autoencoders - arXiv, accessed June 28, 2025, https://arxiv.org/html/2502.08686v1
54. ncclabsustech/DeepSeparator: Deep learning model for EEG artifact removal - GitHub, accessed June 28, 2025, https://github.com/ncclabsustech/DeepSeparator
55. AnEEG: leveraging deep learning for effective artifact removal in EEG data - PubMed, accessed June 28, 2025, https://pubmed.ncbi.nlm.nih.gov/39414897/
56. Can We Ditch Feature Engineering? End-to-End Deep Learning for ..., accessed June 28, 2025, https://www.mdpi.com/1424-8220/20/22/6535
57. Neural Decoding of EEG Signals with Machine Learning: A Systematic Review - PubMed, accessed June 28, 2025, https://pubmed.ncbi.nlm.nih.gov/34827524/
58. (PDF) Can We Ditch Feature Engineering? End-to-End Deep ..., accessed June 28, 2025, https://www.researchgate.net/publication/345976641_Can_We_Ditch_Feature_Engineering_End-to-End_Deep_Learning_for_Affect_Recognition_from_Physiologic

al_Sensor_Data
59. [Literature Review] Toward Fully-End-to-End Listened Speech Decoding from EEG Signals, accessed June 28, 2025, https://www.themoonlight.io/en/review/toward-fully-end-to-end-listened-speech-decoding-from-eeg-signals
60. arxiv.org, accessed June 28, 2025, https://arxiv.org/html/2406.08644v1
61. (PDF) Toward Fully-End-to-End Listened Speech Decoding from EEG Signals, accessed June 28, 2025, https://www.researchgate.net/publication/381404722_Toward_Fully-End-to-End_Listened_Speech_Decoding_from_EEG_Signals
62. "Using AI to decode speech from brain activity" (classification of heard speech using EEG/MEG; first steps towards non-invasive broader speech production) : r/neuralcode - Reddit, accessed June 28, 2025, https://www.reddit.com/r/neuralcode/comments/x5igg3/using_ai_to_decode_speech_from_brain_activity/
63. Brain-to-Voice AI Streams Natural Speech for People with Paralysis - Neuroscience News, accessed June 28, 2025, https://neurosciencenews.com/brain-speech-ai-bci-neurotech-28539/
64. CNN vs. RNN vs. LSTM vs. Transformer: A Comprehensive Comparison | by Emily Smith, accessed June 28, 2025, https://medium.com/@smith.emily2584/cnn-vs-rnn-vs-lstm-vs-transformer-a-comprehensive-comparison-b0eb9fdad4ce
65. Transformer-based neural speech decoding from surface and depth electrode signals - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC11773629/
66. ECA-ATCNet: Efficient EEG Decoding with Spike Integrated Transformer Conversion for BCI Applications - MDPI, accessed June 28, 2025, https://www.mdpi.com/2076-3417/15/4/1894
67. [2501.04359] Decoding EEG Speech Perception with Transformers and VAE-based Data Augmentation - arXiv, accessed June 28, 2025, https://arxiv.org/abs/2501.04359
68. Revolutionary brain-computer interface decoding system - ScienceDaily, accessed June 28, 2025, https://www.sciencedaily.com/releases/2025/03/250327142006.htm
69. Transformer vs RNN: 4 Key Differences and How to Choose - Kolena, accessed June 28, 2025, https://www.kolena.com/guides/transformer-vs-rnn-4-key-differences-and-how-to-choose/
70. CTNet: a convolutional transformer network for EEG-based motor imagery classification, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC11364810/
71. Decoding Covert Speech from EEG Using a Functional Areas Spatio-Temporal Transformer - arXiv, accessed June 28, 2025, https://arxiv.org/html/2504.03762v1
72. What is the difference between acoustic and auditory phonetics? - Quora, accessed June 28, 2025,

https://www.quora.com/What-is-the-difference-between-acoustic-and-auditory-phonetics
73. Speech decoding using cortical and subcortical electrophysiological signals - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2024.1345308/full
74. Brain-to-voice neuroprosthesis restores naturalistic speech - Berkeley Engineering, accessed June 28, 2025, https://engineering.berkeley.edu/news/2025/03/brain-to-voice-neuroprosthesis-restores-naturalistic-speech/
75. Invasive vs. Noninvasive Brain Interface | The ASHA Leader Archive, accessed June 28, 2025, https://leader.pubs.asha.org/doi/10.1044/leader.IN5.18032013.5
76. Few-shot Algorithms for Consistent Neural Decoding (FALCON) Benchmark - bioRxiv, accessed June 28, 2025, https://www.biorxiv.org/content/10.1101/2024.09.15.613126v2.full.pdf
77. Using AI to decode speech from brain activity - Meta AI, accessed June 28, 2025, https://ai.meta.com/blog/ai-speech-brain-activity/
78. The Progress and Prospects of Data Capital for Zero-Shot Deep Brain–Computer Interfaces, accessed June 28, 2025, https://www.mdpi.com/2079-9292/14/3/508
79. Advancements in Zero-Shot and Few-Shot Learning for Large Language Models - GoML, accessed June 28, 2025, https://www.goml.io/blog/advancements-in-zero-shot-and-few-shot-learning-for-large-language-models
80. A zero-shot learning approach to the development of brain-computer interfaces for image retrieval, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC6746355/
81. A Zero-Shot Learning Approach to the Development of Brain-Computer Interfaces for Image Retrieval - PubMed, accessed June 28, 2025, https://pubmed.ncbi.nlm.nih.gov/31525201/
82. (PDF) Few-Shot Learning for EEG Signal Processing: A Model ..., accessed June 28, 2025, https://www.researchgate.net/publication/381888711_Few-Shot_Learning_for_EEG_Signal_Processing_A_Model_Agnostic_Meta_Learning_Approach
83. Harnessing Few-Shot Learning for EEG signal classification: a survey of state-of-the-art techniques and future directions - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2024.1421922/full
84. A Few-Shot Transfer Learning Approach for Motion Intention Decoding from Electroencephalographic Signals - PubMed, accessed June 28, 2025, https://pubmed.ncbi.nlm.nih.gov/38073546/
85. Effects of user mental state on EEG-BCI performance - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC4451337/
86. Effects of Mental Load and Fatigue on Steady-State Evoked Potential Based Brain Computer Interface Tasks: A Comparison of Periodic Flickering and

Motion-Reversal Based Visual Attention | PLOS One, accessed June 28, 2025, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0163426
87. Brain–computer interface in critical care and rehabilitation - PMC - PubMed Central, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC11002623/
88. Eyes on the road: brain computer interfaces and cognitive distraction in traffic - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/journals/neuroergonomics/articles/10.3389/fnrgo.2023.1171910/full
89. Why tracking cognitive load could save doctors and patients, accessed June 28, 2025, https://kevinmd.com/2025/06/why-tracking-cognitive-load-could-save-doctors-and-patients.html
90. Development of a brain-computer interface for patients in the critical care setting | PLOS One, accessed June 28, 2025, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0245540
91. Development of a brain-computer interface for patients in the critical care setting - PubMed, accessed June 28, 2025, https://pubmed.ncbi.nlm.nih.gov/33481888/
92. Brain-computer interfaces and human factors: the role of ... - Frontiers, accessed June 28, 2025, https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2024.1305445/full
93. The Human Factors and Ergonomics of P300-Based Brain-Computer Interfaces - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC4588142/
94. (PDF) The User-Centered Design in Brain-Computer Interface (BCI) research: A novel perspective for evaluating the usability of BCI- controlled applications - ResearchGate, accessed June 28, 2025, https://www.researchgate.net/publication/268804665_The_User-Centered_Design_in_Brain-Computer_Interface_BCI_research_A_novel_perspective_for_evaluating_the_usability_of_BCI-_controlled_applications
95. Evolution of brain-computer interfaces: going beyond classic motor physiology - PMC, accessed June 28, 2025, https://pmc.ncbi.nlm.nih.gov/articles/PMC2920041/
96. "A Non-Invasive Brain-Computer Interface for Real-Time Speech Synthesis" by Jonathan S. Brumberg, Kevin M. Pitt et al. - UNL Digital Commons - University of Nebraska–Lincoln, accessed June 28, 2025, https://digitalcommons.unl.edu/specedfacpub/194/
97. Decoding our thoughts to restore speech - myScience.org, accessed June 28, 2025, https://www.myscience.org/news/2025/decoding_our_thoughts_to_restore_speech-2025-unige
98. The BciAi4SLA Project: Towards a User-Centered BCI - MDPI, accessed June 28, 2025, https://www.mdpi.com/2079-9292/12/5/1234
99. 20 Potential First-Mover Use Cases For Brain-Computer Interface Tech - Forbes, accessed June 28, 2025, https://www.forbes.com/councils/forbestechcouncil/2025/06/27/20-potential-first

-mover-use-cases-for-brain-computer-interface-tech/
100. Accuracy Benchmarking - Speechmatics, accessed June 28, 2025, https://docs.speechmatics.com/tutorials/accuracy-benchmarking
101. Enhancing AI Models: Understanding the Word Error Rate Metric - Galileo AI, accessed June 28, 2025, https://www.galileo.ai/blog/word-error-rate%20-metric
102. Brain-to-Text Benchmark '24: Lessons Learned - arXiv, accessed June 28, 2025, https://arxiv.org/html/2412.17227v1
103. The ethical significance of user-control in AI-driven speech-BCIs: a ..., accessed June 28, 2025, https://www.frontiersin.org/journals/human-neuroscience/articles/10.3389/fnhum.2024.1420334/full
104. BRAIN-COMPUTER INTERFACES: MENTAL PRIVACY, DATA OWNERSHIP, AND THE DAWN OF NEURO-RIGHTS - Jus Corpus, accessed June 28, 2025, https://www.juscorpus.com/brain-computer-interfaces-mental-privacy-data-ownership-and-the-dawn-of-neuro-rights/
105. Ethical issues raised by incorporating personalized language models into brain-computer interface communication technologies: a qualitative study of individuals with neurologi - OHSU, accessed June 28, 2025, https://www.ohsu.edu/sites/default/files/2023-01/REKNEW-Ethical-Issues-Raised-By-Incorporating-Personalized-Language-Models-Into-Brain-Computer-Interface-Communication-Technologies.pdf
106. The Rise of Neurotech and the Risks for Our Brain Data: Privacy and Security Challenges, accessed June 28, 2025, https://www.newamerica.org/future-security/reports/the-rise-of-neurotech-and-the-risks-for-our-brain-data/privacy-and-security-challenges/
107. Navigating the legal and ethical landscape of brain-computer interfaces: Insights from Colorado and Minnesota | IAPP, accessed June 28, 2025, https://iapp.org/news/a/navigating-the-legal-and-ethical-landscape-of-brain-computer-interfaces-insights-from-colorado-and-minnesota
108. Informed Consent in BCI: A Comprehensive Guide - Number Analytics, accessed June 28, 2025, https://www.numberanalytics.com/blog/informed-consent-in-bci-a-comprehensive-guide
109. Translation Ethics Guidelines for Brain-Computer Interface Research - CSET, accessed June 28, 2025, https://cset.georgetown.edu/wp-content/uploads/t0584_brain_computer_ethics_EN.pdf
110. Neural Networks A note on ethical aspects of BCI, accessed June 28, 2025, https://www.neurotechcenter.org/sites/default/files/misc/A%20note%20on%20ethical%20aspects%20of%20BCI.pdf
111. Informed Consent Form Official Study Title: Mind-body Awareness Training and Brain - ClinicalTrials.gov, accessed June 28, 2025, https://cdn.clinicaltrials.gov/large-docs/74/NCT03221374/ICF_002.pdf
112. Family Engagement and Empowerment - ICUDelirium.org, accessed June 28, 2025,

https://www.icudelirium.org/medical-professionals/family-engagement-and-empowerment
113. Brain-computer interfaces for communication, accessed June 28, 2025, https://www.med.upenn.edu/timm/assets/user-content/documents/Fall%202023/Brain-computerinterfacesforcommunication.pdf
114. The Assistive Technology Assessment Process in the ... - OCALI, accessed June 28, 2025, https://ocali.org/up_doc/AT_Resource_Guide_5.pdf
115. Improving the Patient and Family Experience | AHA Trustee Services, accessed June 28, 2025, https://trustees.aha.org/improving-patient-and-family-experience
116. Assistive Technology Consideration Guide - CalTAN.info, accessed June 28, 2025, https://caltan.info/resources/documents/assistive-technology-consideration-guide.pdf
117. Guide to Patient and Family Engagement in Hospital Quality and Safety, accessed June 28, 2025, https://www.ahrq.gov/patient-safety/patients-families/engagingfamilies/index.html