Language brain-computer interfaces: Toward clinically scalable speech neuroprosthesis

Wanwan Guo,[1,2,3] Heng Jiang,[1,2] Haizhou Liu,[1,2] Yuxi Jin,[1,2] Wenjian Qin,[1,2] Hairong Zheng,[1,2] Dong Liang,[1,2] Lingyan Zhang,[3,4,5] and Zhanli Hu[1,2,*]

[1]Research Center for Medical AI, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[2]Key Laboratory of Biomedical Imaging Science and System, Chinese Academy of Sciences, Shenzhen 518055, China
[3]Lab of Molecular Imaging and Medical Intelligence, Department of Radiology, Longgang Central Hospital of Shenzhen, Shenzhen 518116, China
[4]Shenzhen Clinical Medical College, Guangzhou University of Chinese Medicine, Shenzhen 518116, China
[5]Longgang Clinical Institute of Shantou University Medical College, Shenzhen 518116, China
*Correspondence: zl.hu@siat.ac.cn

Received: January 23, 2026; Accepted: April 3, 2026; Published Online: April 17, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100043

Citation: Guo W., Jiang H., Liu H., et al. (2026). Language brain-computer interfaces: Toward clinically scalable speech neuroprosthesis. *The Innovation Informatics* **2**:100043.

Human speech production typically occurs at rates of around 150-200 words per minute (WPM), underscoring the substantial gap between natural communication and artificial interfaces. Recent advances in neural recording technologies and artificial intelligence (AI)-based decoding algorithms have driven a rapid evolution of language brain-computer interfaces (BCIs). A key focus has been on addressing the long-standing trade-off between decoding speed and accuracy, which has historically constrained BCIs to slow, effortful, and cognitively demanding communication.

DEVELOPMENTS ON SPEECH NEUROPROSTHESIS

The year 2023 marked a pivotal milestone in neural speech decoding, with the language decoding rate increasing substantially from 33 to 78 WPM.[1] It has been shown that high-density microelectrode arrays, implanted in speech-related cortical regions, can support neural speech decoding at 62 WPM.[1] This approaches the temporal dynamics of natural language production. Recent advances have further imporved word decoding rate by integrating multimodal information.[1] This included linguistic content, speech acoustics, and facial movements, alongside high-density electrocorticographic recordings. They achieved word decoding rates of up to 78 WPM with an error rate of approximately 25%. Together, these studies represent a qualitative advance in the field. Neural speech decoding has transitioned from generating slow symbolic outputs to closely resembling continuous, conversational communication.

The clinical impact of these advances is clear when compared to conventional augmentative and alternative communication (AAC) systems, which are devices that help people communicate when they cannot rely on natural communicate. Most AAC devices rely on residual motor functions, meaning the user must physically move parts of their body, and involve low-bandwidth, effortful interaction paradigms. Consequently, achievable communication rates remain far below those required for natural conversation, substantially limiting expressive capacity and quality of life for those with profound speech impairments. Speech neuroprostheses, devices that interpret brain signals to produce speech, offer a fundamental paradigm shift rather than just incremental improvement. They aim to restore fluent and natural language by directly decoding neural activity. The workflow of invasive speech neuroprosthesis is shown in Figure 1.

Recent clinical demonstrations reinforced this progress. Card and colleagues implanted four microelectrode arrays into the motor cortex of a 45-year-old individual with amyotrophic lateral sclerosis (ALS). They combined dense intracortical recordings with advanced machine learning to allow rapid calibration and high decoding accuracy.[2] After eight months of training, the system achieved a decoding accuracy of 97.5% with speech output rates approaching 32 WPM. Building on this work, Littlejohn et al. reported a system that continuously transforms cortical activity into natural, intelligible speech in real time.[3] By reducing the latency between speech intention and acoustic output to approximately 80 milliseconds, this system provided a near-natural conversational experience. Notably, to prevent unauthorized decoding of personal inner speech, Kunz and colleagues proposed high-fidelity protection strategies that provide important neural foundations and safeguards for developing safer, more efficient BCIs.[4] Together, these studies provide compelling evidence that fluent, natural speech via invasive BCIs is technically achievable.

Concurrently, these advances highlight a key challenge in the field. All of the language BCIs discussed above use devices that must be implanted in the brain through surgery and require ongoing intracranial recordings. While these methods currently provide unmatched signal fidelity, they also pose substantial barriers to widespread use, including surgical risks, long-term signal stability concern, and limited scalability. As language BCIs transition from proof-of-concept demonstrations to real-world applications, the main question may no longer be whether fluent neural speech decoding is achievable, but rather how these capabilities can be translated into clinically scalable and broadly accessible solutions.

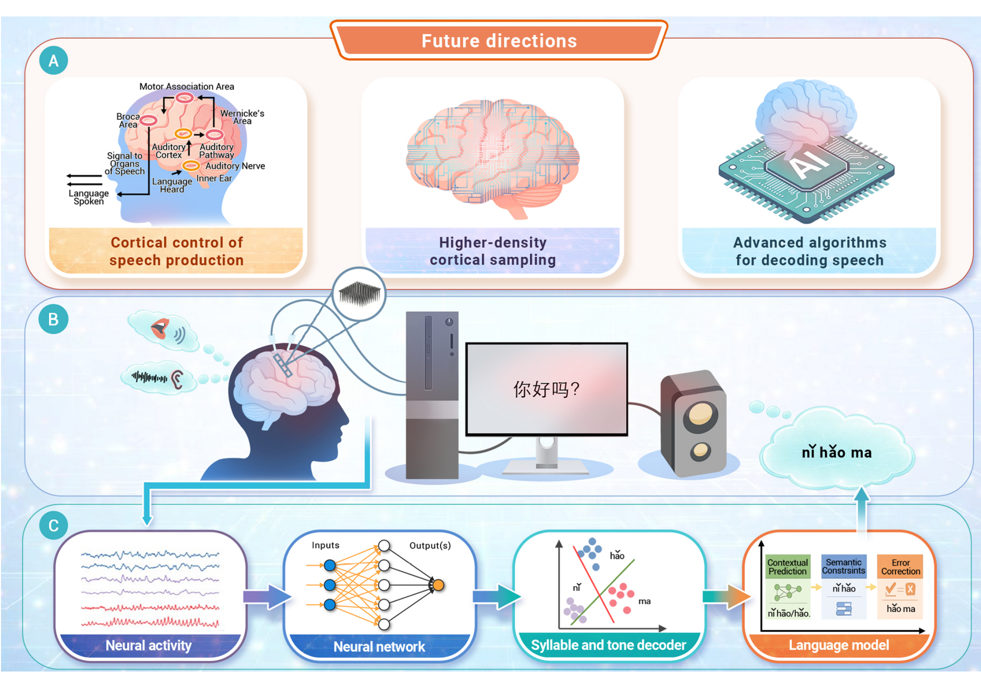

Figure 1. Workflow and future directions of invasive speech neuroprosthesis The illustration presents future directions (A) for brain-to-speech decoding, including enabling precise, real-time cortical control of speech production, employing higher-density cortical sampling to enhance signal fidelity, and developing advanced algorithms capable of decoding naturalistic speech patterns. It further depicts a processing pipeline that begins with invasive brain signals (B) (e.g., high-density microelectrode arrays) and progresses through AI-based language models (C) to achieve speech synthesis.

Current high-performance language BCIs are primarily designed for patients with intact cortical language function but severe motor impairment. Such patients include those with late-stage ALS, brainstem stroke, or lock-in syndrome. These individuals currently rely on AAC systems like eye-tracking and switch-based interfaces. However, these systems remain slow, effortful, and vulnerable to fatigue or disease progression. In contrast, direct neural speech decoding by language BCIs offers the potential for higher-bandwidth, more natural communication. As decoding reliability and surgical safety improve, additional populations with severe motor speech disorders (e.g., individuals with late-stage Parkinson's disease) may benefit. Ultimately, this would help address a substantial unmet clinical need and strengthen the case for clinically scalable deployment.

DEVELOPMENTS ON CHINESE BCIS

A critical yet frequently overlooked challenge in the field of BCIs is linguistic diversity. So far, high-performance language BCIs have been predominantly optimized for Indo-European languages, such as English and Spanish. However, it is still unclear whether decoding strategies developed for these languages can generalize to fundamentally different linguistic systems. Chinese, for example, is a tonal language with a mostly monosyllabic structure and many homophones, imposing unique constraints on both neural encoding and decoding architectures.

Wu and colleagues at Huashan Hospital, Fudan University, have recently addressed this gap by systematically investigating the neural substates of Chinese speech production. Their findings revealed language-specific neural representations which conventional phoneme-based decoding frameworks poorly capture. In 2025, this team reported the first real-time, full-spectrum neural decoding of Chinese speech.[5] To address the high accuracy demands from tonal contrasts and extensive homophony, Qian et al. implemented a syllable-based decoding framework. Using high-density electrocorticographic (ECoG) recordings, they classified 394 distinct Chinese syllables with high-precision.[5] This demonstrates the feasibility of large-vocabulary neural speech decoding in a tonal language. Therefore, this work demonstrates the feasibility of Chinese BCIs and highlights a broader consideration: decoding strategies may need to be tailored to the structural and phonological features of each language, rather than assuming models for Indo-European languages will universally generalize across linguistic systems.

From a clinical perspective, this consideration has profound implications. Speech neuroprostheses aim to restore meaningful communication within each patient's linguistic and sociocultural context. If decoding architectures fail to capture the essential structural features of a given language-such as tonal distinctions or syllabic organization-communication accuracy and usability may be compromised in real-world settings. Therefore, cross-linguistic validity is not merely a theoretical concern but a practical requirement that determines accessibility, equity, and clinical scalability. Robust adaptability across diverse language is essential for translating these transformative technologies from experimental proofs-of-concept to truly global clinical practice.

CURRENT CHALLENGES: Technical, Algorithmic, and Ethical Considerations

Notwithstanding this rapid progress, several critical challenges remain. From a technical standpoint, enhancing the long-term biostability and signal reliability of neural interfaces-particularly for chronic intracortical implants-is a fundamental prerequisite for sustained clinical scalability. In terms of algorithmic development, contemporary decoding models often require extensive subject-specific training data, leading to protracted calibration periods that limit broader scalability. Developing more robust and generalizable models with reduced training demands will be essential to lowering the barriers to widespread clinical deployment. Simultaneously, advances in language BCIs raise significant concerns about data privacy, individual autonomy, and the potential misuse of neural information. Addressing these challenges will require coordinated, multidisciplinary efforts involving clinicians, engineers, ethicists, and policymakers.

FUTURE DIRECTIONS

Moving forward, the evolution of language BCIs will depend on the synergistic advancement of neural encoding and decoding theories, engineering innovation, and AI algorithms, as illustrated in Figure 1.

From a neural encoding and decoding perspective, elucidating the cortical mechanisms underlying speech control-specifically how the sensorimotor cortex coordinates articulation and how the supplementary motor cortex initiates vocalization-will be instrumental in optimizing electrode placement for neuroprosthetic systems.

From an engineering standpoint, iterative improvements in sampling density, signal efficiency, biocompatibility, and the long-term biostability of both surface and penetrating microelectrode arrays remain critical for translating these technologies into routine clinical practice. At the same time, the development of non-invasive and minimally invasive modalities, particularly those integrating emerging sensing technologies such as wearable magnetoencephalography (MEG) with advanced AI algorithms, could help more people use assistive communication beyond specialized clinical settings.

At the computational level, neural speech signals are inherently high-dimensional, non-stationary, and stochastic, presenting substantial challenges for real-time decoding. Modern architectures, including self-supervised learning (SSL), large language models (LLMs), and low-latency machine learning, provide powerful tools for capturing behaviorally relevant latent neural representations and enhancing decoding performance. Equally important is the expansion of cross-linguistic and cross-cultural research, which is indispensable for ensuring that language BCIs evolve into globally inclusive clinical technologies rather than remaining restricted to a narrow subset of languages and experimental contexts.

In summary, language BCIs are redefining the boundaries of human communication. Future research should prioritize refining non-invasive or minimally invasive modalities to expand accessibility beyond specialized environments. Concurrently, building upon recent research in syllable-based decoding for tonal languages, with a particular focus on Chinese neural decoding, is essential to overcome the inherent limitations of phoneme-based models derived from Indo-European languages. Ultimately, clinical scalability will depend not only on decoding accuracy but also on surgical feasibility, long-term stability, cost reduction, and equitable cross-linguistic deployment.

REFERENCES

1. Silva A.B., Littlejohn K.T., Liu J.R., et al. (2024). The speech neuroprosthesis. *Nat. Rev. Neurosci.* **25**:473–492. DOI:10.1038/s41583-024-00819-9
2. Card N.S., Wairagkar M., Iacobacci C., et al. (2024). An accurate and rapidly calibrating speech neuroprosthesis. *N. Engl. J. Med.* **391**:609–618. DOI:10.1056/NEJMoa2314132
3. Littlejohn K.T., Cho C.J., Liu J.R., et al. (2025). A streaming brain-to-voice neuroprosthesis to restore naturalistic communication. *Nat. Neurosci.* **28**:902–912. DOI:10.1038/s41593-025-01905-6
4. Kunz E.M., Abramovich Krasa B., Kamdar F., et al. (2025). Inner speech in motor cortex and implications for speech neuroprostheses. *Cell* **188**:4658–4673. DOI:10.1016/j.cell.2025.06.015
5. Qian Y., Liu C., Yu P., et al. (2025). Real-time decoding of full-spectrum Chinese using brain-computer interface. *Sci. Adv.* **11**:eadz9968. DOI:10.1126/sciadv.adz9968

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of Guangdong Province in China (2023B1515120007, 2024B1515040018), the Shenzhen Science and Technology Program of China (KJZD20240903101307010). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

Hairong Zheng is the Editor-in-Chief of The Innovation Informatics and was blinded from reviewing or making final decisions on the manuscript. Peer review was handled independently of this member and their research group. The other authors declare no conflicts of interest.