Non-invasive brain-to-text decoding: Towards more natural brain-computer interaction

Che Liu,[1,2,4] Xiaoyu Chen,[1,3,4] Changde Du,[1,3] and Huiguang He[1,2,3,*]

[1]State Key Laboratory of Brain Cognition and Brain-inspired Intelligence Technology, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[2]School of Future Technology, University of Chinese Academy of Sciences, Beijing 100049, China
[3]School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China
[4]These authors contributed equally
*Correspondence: huiguang.he@ia.ac.cn

Received: December 11, 2025; Accepted: December 13, 2025; Published Online: January 6, 2026; https://doi.org/10.59717/j.xinn-inform.2026.100021

Citation: Liu C., Chen X., Du C., et al. (2026). Non-invasive brain-to-text decoding: Towards more natural brain-computer interaction. *The Innovation Informatics* 2:100021.

RECENT ADVANCES

Recent advancements in brain-to-text decoding technologies by Meta AI have marked a substantial step forward in non-invasive brain–machine interface (BCI) development. The latest study involving 723 participants demonstrates robust word-level decoding from electroencephalography (EEG) and magnetoencephalography (MEG) across multiple languages in both listening and reading tasks.[1] MEG and reading yield better decoding performance than EEG and listening, with accuracy improving as the amount of training data increases. Notably, the framework exhibits zero-shot decoding capabilities, highlighting its scalability and generalization potential. These findings set a new empirical benchmark for non-invasive brain-to-text systems.

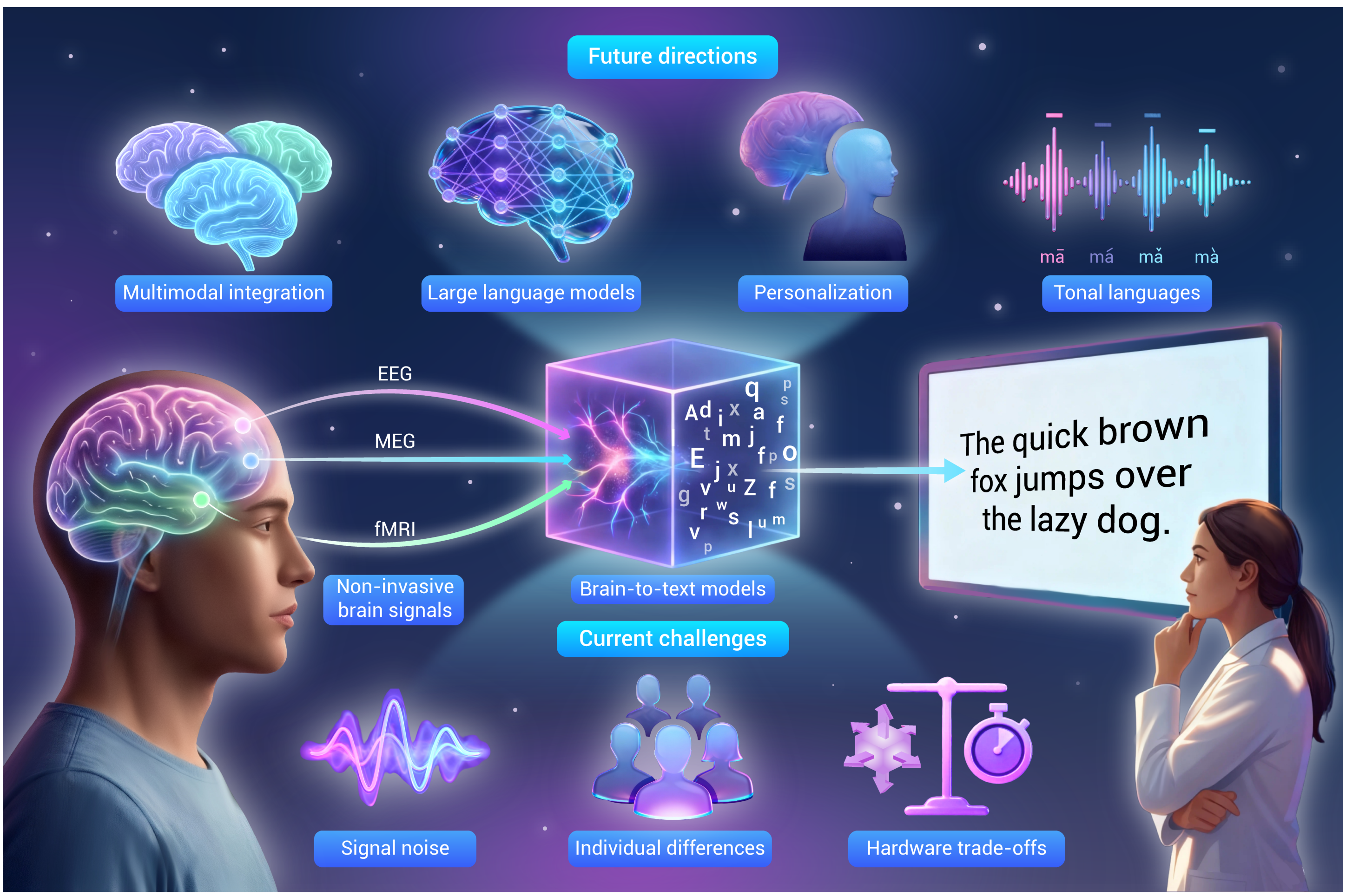

Figure 1. Overview of non-invasive brain-to-text decoding, current challenges, and future directions The illustration presents a pipeline for decoding human language from non-invasive brain signals (EEG, MEG, and fMRI) through brain-to-text models. Current challenges include signal noise, individual differences, and hardware trade-offs. Future directions highlight multimodal integration, large language models, personalization, and tonal language decoding toward more natural and scalable brain–computer interaction.

Prior to this, Meta produced two earlier studies that both used an M/EEG dataset collected during typing tasks, where participants were instructed to type sentences on a QWERTY keyboard. One study introduced a novel Brain2Qwerty model,[2] which decodes sentences directly from EEG and MEG. The model achieved an average character error rate (CER) of just 32% on MEG, significantly outperforming EEG (CER of 67%). This framework offers a potential communication solution for patients with speech impairments and serves as an alternative and complementary paradigm to listening and reading.

The other study focused on the hierarchical neural dynamics underlying language production during typing tasks.[3] Analysis of EEG and MEG revealed sequential activation patterns spanning context, words, syllables, and letters. This hierarchical activation pattern not only validates existing models of linguistic neural coordination but also provides novel insights into cortical information processing during speech formulation. Notably, the discovery of dynamic neural encoding mechanisms demonstrates the brain's capacity for parallel representation of linguistic elements across distinct neural subspaces, effectively preventing informational cross-talk.

Taken together, these studies represent a substantial step forward, showing that M/EEG signals can support reliable text decoding across both perception and production tasks. They also suggest a path toward more practical, scalable and cognitively grounded brain-to-text interfaces, with clear potential for clinical translation in restoring communication for individuals affected by traumatic brain injury or neurodegenerative disorders.

OTHER DEVELOPMENTS

As illustrated in Figure 1, non-invasive brain-to-text decoding refers to the process of decoding human language from non-invasive brain signals using brain-to-text models. In addition to EEG and MEG, other non-invasive brain signals can also be used for decoding language, including functional magnetic resonance imaging (fMRI). For example, Tang et al. decoded continuous language from fMRI in listening tasks.[4] The approach first trained a regression model mapping language representations to fMRI, followed by generating reconstructed sequences word by word using beam search. Experiments showed that the method achieved a BERTScore of 0.81 in speech perception tasks and can further generalize to imagined speech and silent videos. Building on this direction, Chen et al. introduced BP-GPT,[5] a brain-to-text model that further improved decoding performance. The model first used contrastive learning to optimize an fMRI prompt aligned with text representations, and then employed a pre-trained GPT-2 to directly generate sentences. Results showed an average improvement of 4.61% in METEOR and 2.43% in BERTScore.

CURRENT CHALLENGES

Despite notable progress in non-invasive brain-computer interfaces (BCIs), several challenges persist across three key areas: algorithmic, hardware, and data-related.

At the algorithmic level, signal noise and cross-subject variability are significant obstacles. While current methods perform well in controlled laboratory settings, they often struggle in complex, real-world environments where signal noise is more prevalent. This discrepancy arises because laboratory conditions are optimized for signal clarity, whereas real-world scenarios introduce variables such as movement, ambient noise, and varying electrode placements, all of which can degrade signal quality. Additionally, cross-subject variability—differences in individual brain anatomy, neural activity patterns, and cognitive states—complicates the development of universal decoding algorithms. These variations can lead to inconsistent performance across different users, making it challenging to create algorithms that are both accurate and adaptable to a wide range of individuals.

In terms of hardware, achieving optimal temporal and spatial resolution remains a challenge. Different non-invasive devices offer varying advantages and disadvantages in these areas, making it difficult to achieve both simultaneously. For instance, EEG provides excellent temporal resolution, capturing rapid neural dynamics, but its spatial resolution is limited due to the skull's interference with signal transmission. Conversely, fMRI offers high spatial resolution, allowing for precise localization of brain activity, but it lacks the temporal precision needed to track fast neural events. This trade-off necessitates the development of hybrid systems that can integrate multiple modalities to provide comprehensive neural information. Moreover, the practicality of non-invasive BCIs for everyday use is another concern. Many existing devices are bulky, require extensive setup, or are not user-friendly, hindering their adoption outside of clinical or research settings. Advancements in miniaturization, wireless technology, and user-centric design are essential to make these devices more accessible and convenient for daily use.

At the data level, obtaining large amounts of paired brain-language data is a significant issue, especially given the rapid development of large language models. The scale and quality of datasets largely limit the application of large language models in the field of non-invasive BCIs for text generation. Training effective language models requires extensive datasets that capture the nuances of human language and neural representations. However, collecting such data is challenging due to privacy concerns, the need for precise synchronization between neural recordings and linguistic outputs, and the variability in individual neural patterns. Additionally, the lack of standardized protocols for data collection and annotation further complicates the creation of large, high-quality datasets.

FUTURE DIRECTIONS

There are many promising future directions in non-invasive brain-to-text decoding, with the most exciting being hybrid models that integrate multiple BCI techniques and the combination with large language models.

Integrating multiple BCI techniques, such as combining M/EEG with fMRI, offers a comprehensive approach to understanding neural activity. M/EEG provides high temporal resolution, capturing rapid neural dynamics, while fMRI offers high spatial resolution, allowing for precise localization of brain activity. This multimodal integration can enhance the accuracy and reliability of decoding algorithms by leveraging the complementary strengths of each modality.

Moreover, the creation of extensive, high-quality datasets is crucial for advancing non-invasive brain-to-text decoding. Such datasets enable the training of large language models, which can significantly improve the accuracy and efficiency of text generation from brain signals. Collaborative efforts are underway to compile multimodal datasets that combine M/EEG and fMRI data, providing a richer basis for model training.

Based on the high-quality, large-scale datasets, developing pre-trained models, especially large language models, on the datasets and subsequently personalizing them for individual users is a promising direction. By employing continual learning techniques, these models can adapt to the unique neural patterns of each user, effectively addressing cross-subject variability. This approach allows for the creation of robust models that maintain high performance across diverse individuals, facilitating the widespread application of brain-to-text decoding technologies.

Finally, future research should address the unique challenges of decoding tonal languages, such as Mandarin, which exhibit complex linguistic features including tonal variations, syllable structures, and logographic writing systems. Current brain-to-text decoding systems, primarily developed for non-tonal languages like English, struggle to accurately capture tonal nuances and semantic richness in brain signals. Developing specialized models to extract high-precision linguistic information, particularly for tonal and logographic features, remains a critical challenge for advancing brain-to-text decoding in tonal languages.

REFERENCES

1. d'Ascoli S., Bel C., Rapin J., et al. (2025). Towards decoding individual words from non-invasive brain recordings. *Nat. Commun.* **16**:10521. DOI:10.1038/s41467-025-65499-0
2. Lévy J., Zhang M., Pinet S., et al. (2025). Brain-to-text decoding: A non-invasive approach via typing. *Meta Res.* preprint. DOI:10.48550/arXiv.2502.17480
3. Zhang M., Lévy J., d'Ascoli S., et al. (2025). From thought to action: How a hierarchy of neural dynamics supports language production. *Meta Res.* preprint. DOI:10.48550/arXiv.2502.07429
4. Tang J., LeBel A., Jain S., et al. (2023). Semantic reconstruction of continuous language from non-invasive brain recordings. *Nat. Neurosci.* **26**:858–866. DOI:10.1038/s41593-023-01304-9
5. Chen X., Du C., Liu C., et al. (2025). BP-GPT: Auditory neural decoding using fMRI-prompted LLM. *IEEE Int. Conf. Acoust., Speech, Signal Process.* **2025**:1–5. DOI:10.1109/icassp49660.2025.10890142

FUNDING AND ACKNOWLEDGMENTS

This work was supported in part by the Scientific and Technological Innovation (STI) 2030–Major Projects under Grant 2021ZD0201503 and the Lingang Laboratory under Grant No. LGL-1987-07. 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

The authors declare no competing interests.