DEEPER: Dense Electroencephalography Passage Retrieval

Niall McGuire
niall.mcguire@strath.ac.uk
University of Strathclyde
Glasgow, United Kingdom

Yashar Moshfeghi
yashar.moshfeghi@strath.ac.uk
University of Strathclyde
Glasgow, United Kingdom

ABSTRACT

Information retrieval systems have historically relied on explicit query formulation, requiring users to translate their information needs into text. This process is particularly disruptive during reading tasks, where users must interrupt their natural flow to formulate queries. We present *DEEPER* (Dense Electroencephalography Passage Retrieval), a novel framework that enables direct retrieval of relevant passages from users' neural signals during naturalistic reading without intermediate text translation. Building on dense retrieval architectures, *DEEPER* employs a dual-encoder approach with specialised components for processing neural data, mapping EEG signals and text passages into a shared semantic space. Through careful architecture design and cross-modal negative sampling strategies, our model learns to align neural patterns with their corresponding textual content. Experimental results on the ZuCo dataset demonstrate that direct brain-to-passage retrieval significantly outperforms current EEG-to-text baselines, achieving a 571% improvement in Precision@1. Our ablation studies reveal that the model successfully learns aligned representations between EEG and text modalities (0.29 cosine similarity), while our hard negative sampling strategy contributes to overall performance increases.

CCS CONCEPTS

• **Information systems** → *Search interfaces*; **Information retrieval**; *Users and interactive retrieval*;

KEYWORDS

Information Retrieval, EEG, NeuraSearch

ACM Reference Format:
Niall McGuire and Yashar Moshfeghi. 2024. DEEPER: Dense Electroencephalography Passage Retrieval. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 11 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Information Retrieval (IR) systems have historically been conceptualised through the lens of cognitive and behavioural models that characterise how users interact with information. Taylor's fundamental model of information seeking [44] describes how users progress from a visceral, unexpressed information need to a compromised need that can be presented to an information system. Similarly, Kuhlthau's Information Search Process (ISP) model [25] emphasises the cognitive and affective aspects of search, detailing how users move from initial uncertainty and vague thoughts to clearer, more focused understanding through interaction with information systems. These theoretical frameworks highlight a critical challenge in IR: there exists a fundamental gap between users' internal cognitive states and their ability to externalise these states through traditional interaction mechanisms [2, 25, 44]. This "semantic gap" is particularly evident during reading tasks, where users frequently encounter information that prompts new search needs [25]. Under current paradigms, users must interrupt their reading flow to formulate explicit textual queries [2], a process that both disrupts concentration and requires translation of their emerging information needs into concrete search terms. Traditional IR systems, which rely on explicit user interactions through keyboards, mice, or voice commands, forcing users to undergo this translation process, potentially losing critical information about their true needs and intentions.

The emergence of Brain-Computer Interfaces (BCIs) represents a promising direction for enabling more natural, continuous search experiences by allowing systems to detect and respond to users' information needs as they emerge during reading [11, 35]. Beyond enhancing traditional search interactions, BCIs offer critical potential for users with physical impairments who may be unable to express their information needs through conventional input methods, providing a direct neural pathway for information access. Recent neuroscientific studies have demonstrated success in revealing how core IR concepts manifest in the human brain, including the formation of information needs [33, 36], the process of relevance judgment [1, 23], and the experience of satisfaction [40]. Particularly

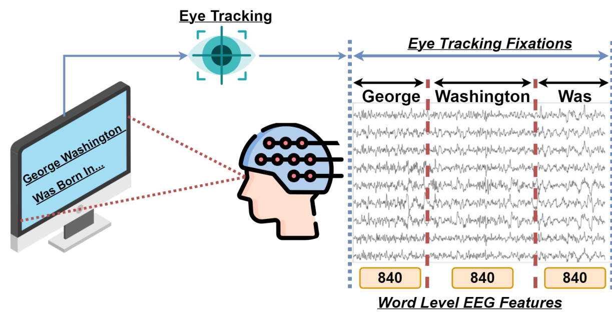


Figure 1: Data collection, illustrating the synchronisation of eye-tracking fixations with EEG recordings during naturalistic reading.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-1-4503-XXXX-X/18/06...$15.00
https://doi.org/XXXXXXX.XXXXXXX

promising are findings showing that neural signatures of relevance and information need emerge in brain activity before users can consciously articulate these judgments [11], suggesting the potential for systems that can proactively identify and retrieve relevant information during natural reading tasks. Several approaches have been explored for neural query formulation, with Steady-State Visually Evoked Potentials (SSVEP) being one of the earliest [7, 46]. SSVEP-based systems attempt to translate brain signals into queries by having users select characters on a virtual keyboard, with each key assigned a unique flicker frequency that can be detected in the user's visual cortex. While innovative, these approaches still require users to interrupt their natural reading flow to explicitly construct queries through character selection, thereby maintaining the fundamental barrier between organic information seeking and query formulation.

More recent research has explored direct brain-to-text translation for query formulation using functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) data [17, 23, 29, 34, 50]. These studies have shown promising results in decoding semantic information from brain activity directly, validating the theoretical possibility of inferring information needs from neural signals. However, both technologies face severe practical limitations that prevent real-world deployment during reading tasks. fMRI requires subjects to lay still within the confined space of the scanner bore, with even slight head movements potentially corrupting the data [37]. Similarly, MEG systems require specially shielded rooms and restrict participant movement [23]. Beyond these physical constraints, both technologies involve substantial equipment costs and spatial requirements that make them impractical for real-world IR applications.

Electroencephalography (EEG) has emerged as a more practical alternative for capturing neural signals during natural reading, offering higher temporal resolution and greater mobility at a fraction of the cost [4]. Recent work has demonstrated that EEG can effectively detect relevance judgments [1], satisfaction [40], and enable direct recommendation of relevant documents from brain signals during reading [11]. These findings suggest the possibility of using EEG signals captured during natural reading as implicit queries, potentially enabling continuous, non-disruptive search experiences. However, current approaches attempting EEG-to-text translation for query formulation [47] have shown limited success in accurately decoding semantic information from EEG signals [19]. Even if such translation were successful, these approaches introduce an unnecessary intermediate step - converting brain signals to text queries before performing retrieval - which risks information loss during translation. Recent findings in neurophysiological IR have shown that neural signatures of relevance emerge before users can consciously articulate their judgments [11, 33]. This suggests that direct neural feedback could capture relevance judgments both earlier and potentially more accurately than traditional explicit feedback methods. Building on this insight and the practical advantages of EEG during reading tasks, we propose an alternative approach to neural query formulation. Rather than attempting to translate brain signals into text queries, we hypothesise that directly utilising EEG signals recorded during natural reading as query representations will achieve better performance than EEG-to-text translation approaches. To address these limitations, we present *DEEPER* (***D**ense **EE**G **P**assage **R**etrieval*), a framework that explores the direct mapping of EEG signals to passage representations for retrieval tasks. Rather than attempting to translate neural signals into explicit text queries, DEEPER projects EEG patterns into the same semantic space as text passages, treating the neural activity itself as an implicit query representation (see Figure 2b). Building on the dual-encoder architecture of Dense Passage Retrieval (DPR) [22], our framework incorporates specialised components for processing neural data and implements cross-modal negative sampling strategies. Our primary contributions include:

- A novel framework for investigating direct EEG-to-passage retrieval without intermediate text translation, providing insights into the feasibility of neural query representations.
- An adaptation of the Bi-Encoder Neural Retrieval Paradigm to facilitate EEG Processing.
- Extensive empirical evaluation comparing direct neural retrieval against current EEG-to-text baselines using the ZuCo dataset.

Our experimental results demonstrate that direct EEG-to-passage mapping consistently outperforms existing EEG query formulation methods that rely on intermediate text translation. This work presents the first demonstration of direct semantic alignment between EEG signals and document representations, showing that meaningful cross-modal retrieval is possible without intermediate translation steps. These findings establish an important direction in neural information retrieval, opening new avenues for research into brain-computer interfaces for information access.

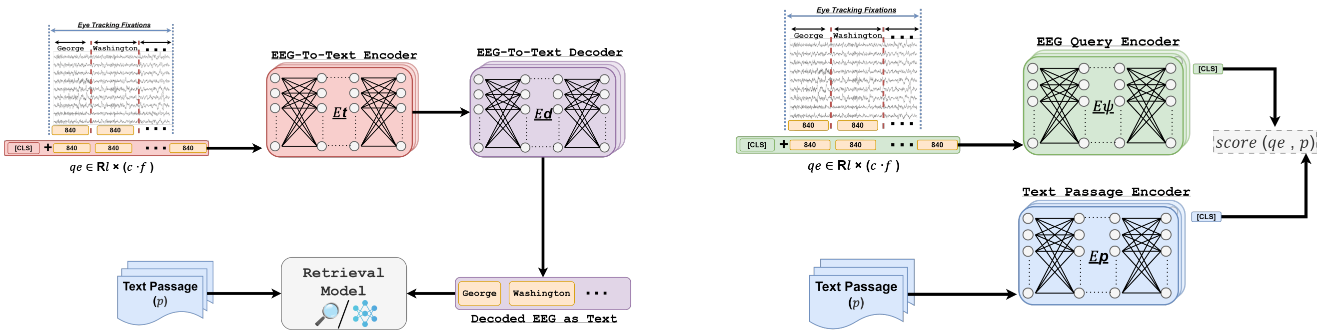
2 RELATED WORKS

2.1 NeuraSearch

Over the past decade, *NeuraSearch* [35] has emerged as a novel interdisciplinary paradigm investigating the neurophysiological basis of information seeking behavior. This field employs non-invasive neuroimaging techniques including magnetoencephalography (MEG) [23], functional magnetic resonance imaging (fMRI) [21, 37], and electroencephalography (EEG) [1, 12, 18, 33] to quantify and characterise the neural processes underlying information retrieval tasks. Empirical investigations have revealed distinct neural signatures associated with key IR processes. Studies of information need formation have identified activation patterns in the anterior cingulate and inferior parietal regions that precede conscious awareness of the need to search [33, 37]. Temporal analysis of relevance judgments has demonstrated neural differentiation between relevant and non-relevant content within 200ms of stimulus presentation [1], with MEG-based decoding achieving 80% classification accuracy [23]. Satisfaction states during information seeking correlate with specific activations in the insula and prefrontal cortex [40], while query formulation processes manifest as modulations in P300 and N400 event-related potentials [21]. These findings establish a quantitative framework for understanding the neural dynamics of information-seeking behaviour.

2.2 Neural Interfaces & IR

The theoretical and empirical foundations established by *NeuraSearch* have enabled increasingly sophisticated brain-computer interface



(a) EEG2Text baseline approach showing the additional decoding step required to translate EEG signals into text queries before retrieval.

(b) DEEPER architecture demonstrating direct matching between EEG signals and text passages.

Figure 2: Overview of competing approaches. (a) Traditional EEG2Text pipeline requires intermediate query decoding before retrieval. (b) Our proposed DEEPER architecture enables direct brain-to-text retrieval without translation.

(BCI) applications in IR systems. Initial work by Eugster et al. [11] demonstrated the feasibility of inferring relevance feedback directly from neural signals during naturalistic reading, enabling implicit document recommendation without explicit user interaction. This breakthrough catalysed further developments in neural feedback mechanisms, with Kauppi et al. [23] achieving robust MEG-based relevance classification and Allegretti et al. [1] implementing real-time EEG-driven result re-ranking. While these feedback-based approaches showed promise in improving search results post-query, researchers recognised the potential for neural signals to enhance earlier stages of the search process. McGuire and Moshfeghi [32] demonstrated this by achieving 90% accuracy in real-time prediction of information need formation using EEG signals during question-answering tasks. Similarly, Ye et al. [50] showed that integrating EEG-based relevance feedback with traditional behavioural signals significantly improves document ranking, particularly when click data is sparse. These successes in interpreting neural signals for both information need detection and result refinement naturally led researchers to explore whether brain activity could be leveraged even earlier—at the query formulation stage itself.

Initial attempts at neural query formulation focused on translating brain signals into text queries through intermediate interfaces, exemplified by Steady-State Visually Evoked Potential (SSVEP) systems. These systems map visual keyboard elements to distinct flicker frequencies, allowing users to input text by focusing on characters that produce corresponding frequency patterns in the visual cortex [46]. While Chen et al. [7] demonstrated a functional SSVEP-based search system achieving 91% accuracy with 5-second average input speeds, these approaches face significant limitations. The requirement for sustained visual attention and character-by-character input not only creates a cognitive burden but also maintains the fundamental constraint of requiring users to consciously translate their information needs into explicit queries. These limitations have motivated investigation into more implicit approaches for neural query formulation, particularly through attempts to decode semantic information directly from brain signals.

2.3 Neural Decoding of Brain Signals

Recent advances in semantic decoding from functional magnetic resonance imaging (fMRI) have demonstrated promising results in reconstructing natural language from brain activity. Ye et al. [50] achieved semantic reconstruction of continuous language by leveraging large language models to map brain representations to semantic space, while Ye et al. [51] demonstrated its potential for query enhancement in retrieval models. MEG has also shown promise as an alternative approach, with studies by Mitchell et al. [34] and Kauppi et al. [23] successfully decoding semantic content with superior temporal resolution. However, both technologies face severe practical limitations that prevent real-world deployment during reading tasks. fMRI requires subjects to lay still within the confined space of the scanner bore, with even slight head movements potentially corrupting the data [37]. Similarly, MEG systems require specially shielded rooms and restrict participant movement [23]. Beyond these physical constraints, both technologies involve substantial equipment costs and spatial requirements that limit their practicality for real-world IR applications.

These constraints have motivated research into decoding language from more accessible recording methods like EEG. The pioneering EEG-to-Text (EEG2Text) work by Wang and Ji [47] framed the decoding task as neural machine translation, using BART to translate EEG signals into text using the ZuCo dataset (see Figure 2a). This approach inspired several works such as DeWave [10], which introduced discrete tokens to improve decoding. However, recent analyses reveal critical methodological limitations in these translation-based approaches. Jo et al. [19] demonstrated that these models rely heavily on teacher forcing during evaluation, resulting in artificially inflated performance metrics. More concerning, these models achieve similar performance when given random noise as input compared to actual EEG signals, suggesting they may primarily memorise training data rather than learn meaningful brain-to-text mappings.

Recent work has attempted to address these limitations by repurposing speech models for EEG decoding. BrainEcho [27] proposed adapting the Whisper model to decode EEG signals directly into text. However, this approach faces inherent constraints when applied to naturalistic reading paradigms, as its audio-centric architecture may not adequately capture reading comprehension dynamics. Furthermore, these methods typically operate with substantially reduced vocabulary sizes compared to datasets like ZuCo (as of writing), limiting their ability to reconstruct complex linguistic content. These limitations, combined with the fundamental challenges in EEG-to-text translation, suggest that attempting to decode explicit queries from neural signals may be an unnecessarily complex intermediate step. Instead, we propose that directly mapping EEG patterns to semantic space could enable more effective neural information retrieval while better preserving the rich information present in brain signals. This insight motivates our investigation into direct EEG-to-passage retrieval, which aims to bridge the gap between users' neural states during reading and their information needs without requiring explicit query formulation.

3 PRELIMINARIES

3.1 Electroencephalography

Electroencephalography (EEG) is a non-invasive neuroimaging method (requiring no surgery or insertion of instruments into the body) that measures and records the electrical activity of the human brain. Within the brain, neurons communicate through electrical impulses, thus generating minute voltage fluctuations on the order of microvolts (millionths of a volt). EEG facilitates the capture and recording of these electrical impulses on a millisecond scale depending on the fixed sampling rate of the system. To capture these signals, sensors known as electrodes are affixed to a participant's scalp. Electrodes are typically comprised of silver/silver chloride (Ag/AgCL), and they detect electrical activity from the underlying cortical areas. The number of electrodes used can vary from one study to another, however, a common setup involves the use of an "EEG cap" - a flexible, lightweight cap with multiple electrodes arranged in standardised positions. This arrangement, known as the "10-20 System", ensures consistency in electrode placement across subjects and studies, allowing for reliable comparison of brain activity from various regions. The electrical signals captured by the electrodes are amplified and digitised for analysis. The sampling rate, measured in Hertz (Hz), determines how many times per second the signal is recorded. For example, a sampling rate of 500 Hz means that the system will capture the electrical activity from the brain 500 times each second.

3.2 Neural Information Retrieval

centred Information retrieval has traditionally centred on matching user information needs to relevant documents through explicit queries. While traditional lexical matching methods like BM25 have proven robust, the emergence of neural retrievers has enabled more sophisticated semantic matching capabilities. These neural approaches typically employ dual-encoder architectures where both queries and documents are mapped to a shared dense vector space, enabling semantic similarity computation. Dense retrievers have shown particular effectiveness in scenarios with sufficient training data, outperforming lexical methods across various IR tasks [22, 24, 42, 49].

A typical neural retriever consists of two main components:

- A query encoder that projects queries into a dense vector space
- A passage encoder that maps documents to the same embedding space

The relevance between a query and passage is typically computed through a similarity function (e.g., cosine similarity, dot product, or L2 distance). Training often employs contrastive learning frameworks that optimise the alignment between queries and their relevant documents while maintaining distance from irrelevant ones.

4 METHODOLOGY

4.1 Task Formulation

While neural retrievers have advanced the state of information retrieval, they still rely on users translating their information needs into explicit textual queries, a process that Belkin [2] identified as inherently lossy and potentially distorting of the original information need. Recent work in brain-computer interfaces has demonstrated the feasibility of decoding semantic information directly from neural signals [5, 17, 34]. While these advances suggest promising directions for enhancing query formulation, current approaches face significant practical and technical limitations when applied to naturalistic reading paradigms [19, 47]. Rather than attempting to translate neural signals into explicit queries, we propose a direct brain-to-passage mapping approach that could better preserve the richness of users' cognitive states during reading while avoiding the complexities of intermediate translation steps.

Formally, let $D = \{w_1, w_2, ..., w_l\}$ be a sequence of l words from a passage being read, and let $E = \{e_1, e_2, ..., e_l\}$ be the corresponding sequence of EEG signals recorded during the reading of each word, where $e_i \in \mathbb{R}^{c \cdot f}$ represents the concatenated features from c EEG channels across f frequency bands for word w_i. Given a large corpus $C = \{p_1, p_2, ..., p_N\}$ of N passages and an EEG signal query representation $q^e \in \mathbb{R}^{l \times (c \cdot f)}$, our goal is to retrieve a small subset of k passages $R_{q^e} = \{p_{i_1}, p_{i_2}, ..., p_{i_k}\} \subset C$ that are most relevant to the reader's information need. This formulation extends traditional dense passage retrieval by replacing text queries with neural signals while maintaining the same passage representation space.

Our model employs a dual-encoder architecture with the following components:

- An EEG encoder $E_\psi : q^e \to \mathbb{R}^d$ that projects brain signals to dense vectors
- A passage encoder $E_P : p \to \mathbb{R}^d$ that maps text to the same embedding space

During inference, relevance between an EEG query and passage is computed as:

$$score(q^e, p) = sim(E_\psi(q^e), E_P(p)) \tag{1}$$

The final retrieval process identifies the top-k most relevant passages by maximising the cumulative relevance scores:

$$R_{q^e} = \underset{R \subset C, |R|=k}{\text{arg, max}} \sum_{p \in R} score(q^e, p) \tag{2}$$

This framework enables direct mapping between neural signals and text passages while preserving the rich information present in brain activity during natural reading tasks.

4.2 EEG Query Encoder

The EEG query encoder architecture builds upon the transformer framework introduced by Vaswani [45], with specific modifications to handle the unique characteristics of neural signals. Given an input EEG sequence $X \in \mathbb{R}^{l \times f}$, where l represents the sequence length and $f = 840$ represents the concatenated features from 105 channels across 8 frequency bands, the encoder first projects the input to the model dimension d through a linear transformation:

$$H_0 = XW_E + b_E \tag{3}$$

where $W_E \in \mathbb{R}^{f \times d}$ and $b_E \in \mathbb{R}^d$ are learnable parameters. A learnable [CLS] token embedding $e_{cls} \in \mathbb{R}^d$ is prepended to the sequence to facilitate global sequence representation:

$$H'_0 = [e_{cls}; H_0] \tag{4}$$

The transformer layers then process this sequence through self-attention and feed-forward networks:

$$H_l = \text{TransformerLayer}_l(H_{l-1}), \quad l \in [1, L] \tag{5}$$

Each transformer layer incorporates multi-head self-attention (MSA) followed by a position-wise feed-forward network:

$$\text{MSA}(Q, K, V) = \text{Concat}(\text{head}_1, ..., \text{head}_h)W^O \tag{6}$$

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \tag{7}$$

Our empirical investigation of sequence pooling strategies revealed that utilising the [CLS] token representation consistently outperforms both mean and max pooling alternatives for EEG sequence representation. This finding aligns with previous work in cross-modal representation learning [28, 30], suggesting that the [CLS] token effectively aggregates global sequence information through self-attention mechanisms. The final EEG query representation is therefore computed as:

$$q = W_O H_L[0] + b_O \tag{8}$$

where $H_L[0]$ represents the [CLS] token from the final layer, and W_O, b_O are learnable projection parameters mapping to the final output dimension.

4.3 EEG Negative Sampling

Effective negative sampling strategies have proven crucial for contrastive learning in dense retrieval systems [22, 49]. Our approach combines in-batch negative sampling with hard negative mining to create a robust training signal. For each EEG query q^e in a batch, we utilise a multi-faceted negative sampling strategy that leverages both in-batch dynamics and semantic similarity. In-batch negative sampling treats all other query-passage pairs (q^e, p^+) within a batch of size B as potential negatives, providing $B - 1$ diverse negative examples while maintaining computational efficiency through batch-wise computation. To handle the complexity of our dataset where multiple subjects read identical passages, we implement a careful negative sampling mechanism using subject-specific passage lookup tables defined as:

$$L(p, s) = \{(p_i, s_j) \in P \times S : p_i \neq p\} \tag{9}$$

where P represents the set of all passages and S the set of all subjects. These lookup tables prevent the incorrect selection of neural responses from different subjects reading the same passage as negative examples, ensuring the model learns genuine semantic distinctions rather than subject-specific variations. For hard negative mining, we exploit the semantic structure of our text corpus. Given a positive passage p^+, we identify semantically similar passages using BM25:

$$H(p^+) = \text{top-k}_{p \in C \setminus \{p^+\}} \text{BM25}(p^+, p) \tag{10}$$

where C represents the corpus and $k = 5$ in our implementation. The corresponding EEG recordings e_{s,p_h} from these hard negative passages $p_h \in H(p^+)$ are incorporated into the negative pool. This approach is particularly effective given neuroscientific evidence suggesting consistent neural patterns for similar semantic content [17, 34]. The training objective unifies these negative sampling strategies in a contrastive learning framework:

$$\mathcal{L} = -\log \frac{\exp(s(q^e, p^+)/\tau)}{\exp(s(q^e, p^+)/\tau) + \sum_{p^- \in N} \exp(s(q^e, p^-)/\tau)} \tag{11}$$

where $N = N_b \cup N_r \cup N_h$ represents the combined set of in-batch (N_b), random (N_r), and hard (N_h) negatives, and τ is the temperature parameter controlling the sharpness of the distribution. This comprehensive negative sampling approach enables the model to learn robust representations that effectively distinguish between neural patterns associated with semantically similar but distinct content, while properly handling the many-to-one relationships inherent in our dataset's structure.

4.4 Dataset Creation

Training neural retrieval models requires pairs of queries and relevant documents, along with carefully selected negative examples. However, unlike traditional IR datasets, EEG datasets of sufficient scale with predefined query-document pairs are not readily available. To address this limitation, we adapt the inverse cloze task (ICT) framework [6, 26] to construct synthetic training data from our EEG recordings. The ICT approach enables us to generate query-document pairs by treating spans of text as implicit queries while considering their surrounding context as relevant documents. This setup is particularly well-suited for our scenario, as it mirrors the natural reading process captured in the ZuCo dataset [15, 16] where EEG signals are recorded as 30-subjects read continuous text passages.

Formally, given a document represented as a sequence of tokens $D = \{w_1, w_2, \ldots, w_m\}$ and its corresponding EEG recordings $E = \{e_1, e_2, \ldots, e_m\}$ where $e_i \in \mathbb{R}^{c \cdot f}$ represents the neural signal captured during the reading of token w_i, we extract a text span $Q = \{w_l, w_{l+1}, \ldots, w_r\}$ to serve as a pseudo-query. The EEG signals corresponding to this span $q^e = \{e_l, e_{l+1}, \ldots, e_r\}$ form our query representation. With probability p_{mask} (set to 0.9 in our implementation), we remove this span from the document to form the positive document: $D \setminus Q = \{w_1, \ldots, w_{l-1}, w_{r+1}, \ldots, w_m\}$. Otherwise, with probability $1 - p_{mask}$, we retain the query span in the document, creating a more challenging learning scenario where the model must learn robust matching strategies beyond exact token matching.

Algorithm 1 Inverse Cloze Test for Neural Query Generation

Require: Document tokens $D = \{w_1, w_2, \ldots, w_m\}$, EEG signals $E = \{e_1, e_2, \ldots, e_m\}$, mask probability p_{mask}
Ensure: EEG query q^e, modified document D'
$L \leftarrow \lfloor m \cdot 0.3 \rfloor$ ▷ Set query length to 30% of document
Select random index l where $0 \leq l \leq m - L$
$Q \leftarrow \{w_l, w_{l+1}, \ldots, w_{l+L-1}\}$ ▷ Extract text span
$q^e \leftarrow \{e_l, e_{l+1}, \ldots, e_{l+L-1}\}$ ▷ Extract corresponding EEG
$u \sim \text{Uniform}(0, 1)$ ▷ Sample uniform random number
if $u < p_{mask}$ **then**
  $D' \leftarrow \{w_1, \ldots, w_{l-1}, w_{l+L}, \ldots, w_m\}$ ▷ Remove span
else
  $D' \leftarrow D$ ▷ Keep original document
end if
return q^e, D'

Algorithm 1 details our ICT implementation. The algorithm takes as input the document tokens, their corresponding EEG signals, query length ratio, and mask probability, then:

(1) Computes the query length as a fraction of the total document length
(2) Randomly selects a starting position for the query span
(3) Extracts both the text query and its corresponding EEG signals
(4) Randomly decides whether to remove the query span from the document based on p_{mask}
(5) Returns the EEG query and document pair for training

Our data processing pipeline incorporates several key design choices that promote robust model learning. First, random span selection prevents the model from learning positional biases or developing heuristics based on span location. This randomisation ensures the model must learn to match content semantically rather than relying on position-based patterns. Second, the probabilistic removal of the query span prevents the model from overly relying on exact token matching, encouraging it to develop more sophisticated semantic matching strategies. Third, the query length is determined by a consistent ratio ($q = 0.3$ in our implementation) of the document length, ensuring queries are proportional to document length while maintaining sufficient context in both query and document.

5 EXPERIMENTAL SETUP

In this work, we aim to address the following research questions:

- **RQ1:** To what extent can EEG signals recorded during naturalistic reading be effectively utilised as implicit queries for passage retrieval without intermediate text translation?
- **RQ2:** How does direct EEG-to-passage retrieval (*DEEPER*) compare with existing EEG-to-text translation approaches in terms of retrieval effectiveness?
- **RQ3:** Can EEG signals be generalised across participants to facilitate passage retrieval?
- **RQ4:** How does the retrieval effectiveness of neural queries compare to traditional text-based retrieval methods?

5.1 Implementation

Our EEG encoder consists of a transformer neural network with 6 layers and 8 attention heads. Each transformer layer contains a multi-head self-attention mechanism followed by layer normalisation and a position-wise feed-forward network. The input dimension of 840 (corresponding to 105 channels × 8 frequency bands) is projected to a model dimension of 512. A dropout rate of 0.3 is applied throughout the network for regularisation. The training was conducted on an NVIDIA H100 GPU. We utilise the AdamW optimiser with an initial learning rate of 1e-4 and linear warm-up over the first 10% of steps, followed by a linear decay schedule. The training batch size is set to 32 with gradient accumulation every 4 steps for an effective batch size of 128. Gradient clipping is employed with a maximum norm of 1.0. Layer normalisation epsilon is set to 1e-6. Model checkpointing occurs at the end of each epoch, saving the weights that achieve the lowest validation loss. Early stopping is implemented with a patience of 5 epochs. For the contrastive learning objective, we employ *InfoNCE* loss [39] with a temperature parameter T = 0.07. During training, we use a mix of in-batch negatives and hard negatives sampled from the top-k retrieved passages that are semantically similar but not temporally aligned with the EEG signals.

Table 1: Dataset statistics and lexical overlap between splits.

Metric	Train	Dev	Test
Number of examples	10,391	1,292	1,367
Total unique queries	783	101	104
Total unique passages	784	101	104
Total words	166,623	19,718	23,216
Unique words	3,716	717	814
Avg. passage length	16.0 ± 7.0	15.3 ± 7.7	17.0 ± 6.9
Avg. query length	6.4 ± 3.1	6.1 ± 3.5	6.8 ± 2.9
Lexical Overlap			
Train-Split	–	0.105	0.115
Dev-Split	–	–	0.148

5.2 Training & Evaluation

To comprehensively evaluate our proposed approach and address our research questions, we implement a multi-faceted experimental framework that examines both the effectiveness of direct neural retrieval and its generalisation capabilities. All experiments are conducted using 5-fold cross-validation, where the dataset is divided into five equal parts, with each fold serving as the test set once while the remaining folds constitute the training data. Within each fold, we further partition the training data to maintain our 80-10-10 split between training, development, and test sets. Our primary evaluation compares three distinct methodological approaches: direct EEG-to-passage retrieval (DEEPER), EEG-to-text translation followed by retrieval (EEG2Text+Retriever), and traditional text-based retrieval as a reference point.

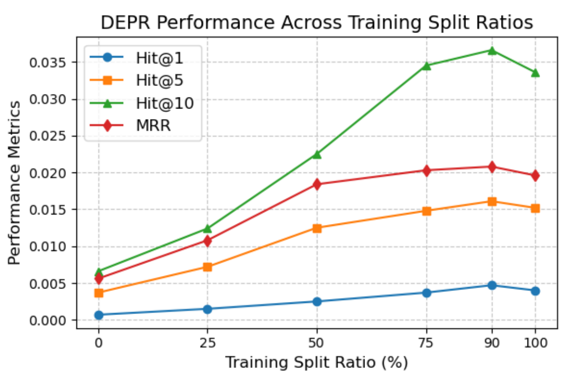
This three-way comparison allows us to assess both the viability of direct neural retrieval (RQ1) and its performance relative

Table 2: Retrieval performance comparison across different query modalities and retrieval methods. Results are averaged across 5-fold cross-validation, with models evaluated on test sets containing varying degrees of query-passage overlap. * indicates statistically significant improvement over random noise baseline using a paired t-test (p < 0.05).

Query Modality	Retriever	Precision@1	Precision@5	Precision@10	MRR
Noise	DEEPER(ours)	.0007 (±.0002)	.0037 (±.0008)	.0066 (±.0012)	.0056 (±.0011)
EEG	E2T+BM25	.0007 (±.0002)	.0032 (±.0007)	.0064 (±.0013)	.0051 (±.0010)
	E2T+ColBERT	.0007 (±.0002)	.0034 (±.0008)	.0062 (±.0012)	.0053 (±.0011)
	DEEPER(ours)	.0047* (±.0009)	.0161* (±.0025)	.0366* (±.0042)	.0206 (±.0031)
Text	BM25	.0128* (±.0221)	.0538* (±.0264)	.0825* (±.0289)	.0531* (±.0268)
	ColBERT	.0361* (±.0341)	.0834* (±.0392)	.1027* (±.0415)	.0784* (±.0387)

to existing approaches (RQ2). We utilise the ZuCo dataset, generating query-passage pairs using our modified ICT framework (Algorithm 1). Through empirical testing of various masking probabilities ranging from 0.0-1.0, we found that $p_{mask} = 0.9$ provides the optimal balance between query span removal and retention during training, maximising the model's ability to learn robust semantic relationships while maintaining sufficient context (see Figure 3).

Figure 3: DEEPER model performance at query-passage overlap ratios



To investigate semantic understanding beyond surface-level patterns (RQ1, RQ4), we construct five distinct test sets with systematically varying degrees of query-document lexical overlap [0%, 25%, 50%, 75%, and 100% text removed]. This evaluation approach enables us to distinguish between models that rely primarily on exact matching versus those that capture deeper semantic relationships. The EEG2Text baseline model follows the implementation parameters specified in Jo et al. [19], with decoded queries evaluated using both BM25 (lexical baseline) and ColBERT (neural retrieval model fine-tuned on our training split).

Following standard practice in neural retrieval evaluation with binary relevance judgments, we report Mean Reciprocal Rank (MRR) and Precision at different cutoffs (Precision@k for k = 1,5,10). These metrics are particularly suitable for our ICT-based evaluation setup, where passages containing the removed span that generated the EEG query are considered relevant, while all other passages are treated as non-relevant. All metrics are computed using 5-fold cross-validation to ensure robust evaluation across different data splits.

Recent work by Jo et al. [19] has raised important methodological concerns about EEG-to-text models, demonstrating that some systems perform similarly whether given real EEG signals or random noise as input. To address these concerns and ensure the validity of our approach (RQ2), we conduct additional control experiments using random noise queries scaled to match the statistical properties (mean, standard deviation, range) of the real EEG signals in our test set. This validation step is crucial for demonstrating that our model learns meaningful neural-semantic mappings rather than exploiting dataset artefacts or statistical regularities.

6 RESULTS

DEEPER enables direct brain-to-passage retrieval. Our results (Table 2) demonstrate that DEEPER successfully enables direct matching between EEG signals and text passages without intermediate translation, achieving statistically significant improvements over random noise queries across all metrics (p < 0.05). The direct brain-to-passage approach shows relative gains of 571% in Precision@1 (.0047 vs .0007) and 454% in Precision@10 (.0366 vs .0066) compared to noise baselines. This substantial improvement across different retrieval cutoffs indicates DEEPER captures meaningful patterns in neural signals rather than exploiting dataset artifacts. The consistent gains at both early and later ranks suggest the model learns robust neural-semantic representations that maintain their discriminative power even for more challenging retrieval scenarios.

DEEPER improves over EEG-To-Text. Both baseline EEG-to-text methods perform nearly identically to random noise queries, with Precision@1 scores of .0007 and MRR around .005 (Table 2). These results suggest that attempting to decode explicit queries from neural signals introduces unnecessary information loss. The failure of both traditional lexical (BM25) and neural (ColBERT) retrievers when operating on decoded queries further validates our hypothesis that intermediate text translation creates a problematic bottleneck. In contrast, DEEPER's direct mapping approach achieves a Precision@1 of .0047 and MRR of .0206, demonstrating that preserving the rich information in neural signals through end-to-end learning leads to substantially better retrieval performance.

Direct EEG retrieval shows promise despite text-EEG gap. While there remains a performance gap between EEG-based and text-based retrieval (BM25 Precision@1: .0128, ColBERT Precision@1: .0361), DEEPER demonstrates that direct brain-to-passage mapping is feasible (Table 2). The relative performance difference between DEEPER and text retrieval (approximately 3x) is notably smaller than the gap between DEEPER and EEG-to-text baselines (approximately 5x), suggesting direct mapping may be a more promising direction for bridging neural and text retrieval. This smaller gap is particularly encouraging given the inherent challenges of working with neural signals compared to clean text input and suggests that further advances in neural signal processing and representation learning could continue to narrow this divide.

DEEPER displays robust performance across evaluation. The stability of DEEPER's performance is evidenced by the relatively low standard deviations across all metrics (approximately 15-20% of mean values) (Table 2). This consistency is particularly noteworthy given both the inherently noisy nature of EEG signals and our challenging evaluation setting: DEEPER maintains statistically significant improvements over baselines across 30 participants engaged in naturalistic reading tasks with unconstrained vocabulary. Unlike previous approaches that often rely on controlled vocabularies or specific task conditions, our model demonstrates robust performance in an open vocabulary setting where readers encounter natural language variations. The model's ability to handle both neural signal variability across participants and linguistic variability in natural reading suggests strong potential for real-world applications. This robustness is further validated by consistent performance across varying query-passage overlap conditions, indicating DEEPER learns meaningful neural-semantic mappings that generalise beyond surface-level patterns.

These results not only validate the feasibility of direct EEG-to-passage retrieval but also establish an important new direction in neural information retrieval. By demonstrating that brain signals can be effectively mapped to passage representations without intermediate translation steps, DEEPER opens new possibilities for more natural and efficient search interactions through brain-computer interfaces.

7 ABLATION

To better understand the effectiveness of our approach, we conduct two key ablation studies: (1) an analysis of learned representation spaces between EEG and text modalities, and (2) an evaluation of our hard negative sampling strategy's impact on retrieval performance.

7.0.1 Cross-Modal Representation Analysis. To understand how effectively our model maps between neural and textual modalities, we first examine the learned representation space through t-SNE visualisation of our test set (Figure 4). The visualisation reveals distinct clustering behaviours between EEG signals (blue) and text passages (red), with six annotated pairs representing the spectrum of cross-modal alignment from weakest (Pair 1, sim: 0.002) to strongest (Pair 6, sim: 0.537). The EEG embeddings demonstrate structured organisation, particularly visible in the curved formation at the right, while text embeddings exhibit a more dispersed distribution with multiple semantic clusters. High-performing pairs (Pairs 4, 5, and 6 with similarities of 0.437, 0.444, and 0.537 respectively) show close spatial proximity between their EEG and text representations, with an average similarity score of 0.29 ± 0.21 across all inputs, suggesting successful semantic alignment in these cases.

Given these encouraging patterns in visualisation, a natural question arises: how does EEG-based retrieval compare to traditional text-based approaches when controlling for architectural and data differences? To investigate this, we train a parallel text query encoder following the same architectural design as our EEG encoder using the same number of samples from the ZuCo dataset, replacing the 840-dimensional EEG input features with 768-dimensional BERT token embeddings. Table 3 shows that our EEG encoder achieves comparable performance to the text encoder across all metrics, with only marginally lower scores (Precision@1: .0047 vs .0058, MRR: .0206 vs .0289). This relatively small performance gap when controlling for architecture and dataset size suggests that neural signals can effectively serve as query representations with nearly the same effectiveness as explicit text queries.

7.0.2 Impact of Hard Negative Sampling. To evaluate the effectiveness of our proposed hard negative sampling strategy, we compare our full model against a variation trained using only random negative sampling. Table 4 presents the results:

The results demonstrate that incorporating hard negatives during training leads to consistent improvements across all metrics,

Table 3: Retrieval performance comparison between Text and EEG query encoders

Modality	Precision@1	Precision@5	Precision@10	MRR
Text	.0058	.0209	.0424	.0289
EEG	.0047	.0161	.0366	.0206

Figure 4: t-SNE embedding space visualisation from EEG query and Text passage encoders

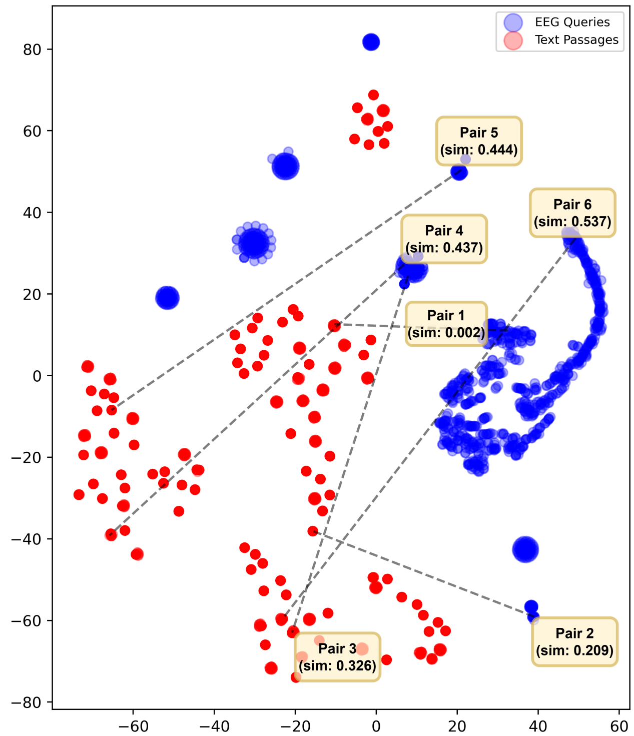


Table 4: Impact of hard negative sampling on retrieval performance

Negative	Precision@1	Precision@5	Precision@10	MRR
DEEPER (Rnd)	.0031	.0142	.0298	.0173
DEEPER (hard)	**.0047**	**.0161**	**.0366**	**.0206**

with relative gains of 51.6% in Precision@1 and 19.1% in MRR. This suggests that exposing the model to challenging negative examples during training helps it develop more robust discriminative features for cross-modal matching. The performance gap is particularly pronounced for Precision@1, indicating that hard negative sampling is especially beneficial for precise retrieval tasks where only the top result matters.

These ablations provide several key insights: (1) despite the inherent noise and variability in EEG signals, our model learns representations that achieve near-parity with text-based queries when controlling for architecture and dataset size, (2) the visualisation reveals clear structure in how neural signals are mapped to semantic space, with strong pairs showing close spatial alignment, and (3) our hard negative sampling strategy plays a crucial role in improving the model's ability to distinguish between semantically similar passages. Together, these findings validate both our architectural choices and training strategy while highlighting promising directions for future improvements in brain-to-text retrieval.

8 DISCUSSION AND CONCLUSION

This work demonstrates that direct mapping between EEG signals and passage representations enables more effective retrieval compared to intermediate translation approaches. The performance differential between DEEPER and EEG-to-text baselines (571% improvement in Precision@1) aligns with findings that preserving high-dimensional neural patterns is crucial for brain-computer interfaces [5, 34]. While traditional text-based methods show stronger absolute performance, similar to patterns observed in other cross-modal tasks [28, 30], our controlled experiments with architecturally identical query encoders trained on matched quantities of data show only modest differences in performance between EEG and text approaches (Precision@1: .0047 vs .0058). This relative performance difference is particularly encouraging given the inherent challenges of working with noisy neural signals compared to clean textual input, suggesting that direct mapping may be a promising direction for bridging neural and text retrieval.

Our findings advance discussions about brain-based information retrieval architectures [11, 35] in several key ways. First, direct brain-to-passage mapping could benefit users during naturalistic reading tasks, where traditional query formulation requires interrupting the reading flow and explicitly translating information needs into text [44]. Moreover, the system offers potential benefits for users with physical impairments who may struggle with traditional text input methods [48], providing a more accessible pathway to information access. Second, our contrastive learning framework's effectiveness in aligning neural and textual representations (cosine similarity 0.29 ± 0.21) extends cross-modal representation learning [28, 30] to neural signals, demonstrating that meaningful semantic alignment between brain activity and text is achievable above random noise. This alignment, achieved in a naturalistic reading environment, suggests that neural patterns during reading contain recoverable semantic information that can be effectively mapped to textual representations. Third, our performance gains through hard negative sampling (51.6% improvement in Precision@1) validate the importance of careful training strategies for cross-modal learning with neural data, particularly in addressing the unique challenges posed by the high dimensionality and noise characteristics of EEG signals.

As a foundational work in direct brain-to-passage retrieval, our approach opens several promising research directions. The single-vector representation approach establishes a baseline while leaving room for more sophisticated modelling approaches such as multi-vector representations that have been demonstrated in both dense retrieval [14, 24] and re-ranking cross-encoder contexts [31, 38]. These methods could enable more nuanced brain-text matching, while EEG encoder pre-training could improve representation quality [9]. The relatively small performance gap between EEG and text encoders when trained on matched data, combined with recent work showing favourable scaling properties of EEG-based models [43], suggests that increased dataset sizes could substantially improve performance. This is particularly important for capturing the wide variability of neural responses across different readers and reading contexts, especially given the current limitations of our ICT-based training approach. Our use of the Inverse Cloze Task, while enabling systematic evaluation, points to opportunities for developing dedicated datasets that capture neural signals during active information seeking with explicit relevance judgments [18]. Beyond naturalistic reading, future work could explore further perceptual modalities including auditory processing [3, 8] and imagined speech [41], which engage distinct neural mechanisms [13, 20]. Understanding how these different processing pathways affect retrieval performance could expand the utility of brain-based search interfaces across diverse interaction scenarios, from multimedia content exploration to hands-free information access [48].

In conclusion, *DEEPER* represents an important step toward improving the accessibility and efficiency of information retrieval systems through direct neural interfaces. By demonstrating direct EEG-to-passage mapping, our work establishes a promising direction for neural IR research that could benefit both users with specific accessibility needs and those seeking more efficient information access methods. As the field advances, direct neural feedback mechanisms could enhance how users interact with search systems, particularly in scenarios where traditional input methods are impractical or limiting. This foundation enables the development of more accessible information systems aligned with human cognitive processes, while our experimental results provide evidence that direct brain-to-passage retrieval is both feasible and potentially beneficial to current IR approaches. The success of *DEEPER* in achieving significant performance improvements over baselines while maintaining robust performance across different evaluation conditions suggests that direct neural retrieval could complement existing IR systems, expanding the range of available interaction methods for diverse user needs and contexts.

REFERENCES

[1] Marco Allegretti, Yashar Moshfeghi, Maria Hadjigeorgieva, Frank E Pollick, Joemon M Jose, and Gabriella Pasi. 2015. When relevance judgement is happening? An EEG-based study. In *Proceedings of the 38th international acm sigir conference on research and development in information retrieval*. 719–722.

[2] Nicholas J Belkin. 1980. Anomalous states of knowledge as a basis for information retrieval. *Canadian journal of information science* 5, 1 (1980), 133–143.

[3] Shohini Bhattasali, Jonathan Brennan, Wen-Ming Luh, Berta Franzluebbers, and John Hale. 2020. The Alice Datasets: fMRI & EEG observations of natural language comprehension. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*. 120–125.

[4] CD Binnie and PF Prior. 1994. Electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry* 57, 11 (1994), 1308–1319.

[5] Charlotte Caucheteux, Alexandre Gramfort, and Jean-Rémi King. 2022. Deep language algorithms predict semantic comprehension from brain activity. *Scientific reports* 12, 1 (2022), 16327.

[6] Wei-Cheng Chang, Felix X Yu, Yin-Wen Chang, Yiming Yang, and Sanjiv Kumar. 2020. Pre-training tasks for embedding-based large-scale retrieval. *arXiv preprint arXiv:2002.03932* (2020).

[7] Xuesong Chen, Ziyi Ye, Xiaohui Xie, Yiqun Liu, Xiaorong Gao, Weihang Su, Shuqi Zhu, Yike Sun, Min Zhang, and Shaoping Ma. 2022. Web search via an efficient and effective brain-machine interface. In *Proceedings of the fifteenth ACM international conference on web search and data mining*. 1569–1572.

[8] Wendy A de Heer, Alexander G Huth, Thomas L Griffiths, Jack L Gallant, and Frédéric E Theunissen. 2017. The hierarchical cortical organization of human speech processing. *Journal of Neuroscience* 37, 27 (2017), 6539–6557.

[9] Jacob Devlin. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[10] Yiqun Duan, Jinzhao Zhou, Zhen Wang, Yu-Kai Wang, and Chin-Teng Lin. 2023. Dewave: Discrete eeg waves encoding for brain dynamics to text translation. *arXiv preprint arXiv:2309.14030* (2023).

[11] Manuel JA Eugster, Tuukka Ruotsalo, Michiel M Spapé, Oswald Barral, Niklas Ravaja, Giulio Jacucci, and Samuel Kaski. 2016. Natural brain-information interfaces: Recommending information by relevance inferred from human brain signals. *Scientific reports* 6, 1 (2016), 38580.

[12] Jacek Gwizdka, Rahilsadat Hosseini, Michael Cole, and Shouyi Wang. 2017. Temporal dynamics of eye-tracking and EEG during reading and relevance decisions. *Journal of the Association for Information Science and Technology* 68, 10 (2017), 2299–2312.

[13] Gregory Hickok and David Poeppel. 2007. The cortical organization of speech processing. *Nature reviews neuroscience* 8, 5 (2007), 393–402.

[14] Sebastian Hofstätter, Markus Zlabinger, and Allan Hanbury. 2020. Interpretable & time-budget-constrained contextualization for re-ranking. In *ECAI 2020*. IOS Press, 513–520.

[15] Nora Hollenstein, Jonathan Rotsztejn, Marius Troendle, Andreas Pedroni, Ce Zhang, and Nicolas Langer. 2018. ZuCo, a simultaneous EEG and eye-tracking resource for natural sentence reading. *Scientific data* 5, 1 (2018), 1–13.

[16] Nora Hollenstein, Marius Troendle, Ce Zhang, and Nicolas Langer. 2019. ZuCo 2.0: A dataset of physiological recordings during natural reading and annotation. *arXiv preprint arXiv:1912.00903* (2019).

[17] Alexander G Huth, Wendy A De Heer, Thomas L Griffiths, Frédéric E Theunissen, and Jack L Gallant. 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532, 7600 (2016), 453–458.

[18] Kaixin Ji, Danula Hettiachchi, Flora D Salim, Falk Scholer, and Damiano Spina. 2024. Characterizing information seeking processes with multiple physiological signals. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1006–1017.

[19] Hyejeong Jo, Yiqian Yang, Juhyeok Han, Yiqun Duan, Hui Xiong, and Won Hee Lee. 2024. Are EEG-to-Text Models Working? *arXiv preprint arXiv:2405.06459* (2024).

[20] Nayereh Joodi and Farzaneh Rahmani. 2019. Application of Functional Magnetic Resonance Imaging in Neurolinguistics: A Systematic Review. *Frontiers in Biomedical Technologies* (2019).

[21] Lauri Kangassalo, Michiel Spapé, Giulio Jacucci, and Tuukka Ruotsalo. 2019. Why do users issue good queries? neural correlates of term specificity. In *Proceedings of the 42nd international acm sigir conference on research and development in information retrieval*. 375–384.

[22] Vladimir Karpukhin, Barlas Oğuz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih. 2020. Dense passage retrieval for open-domain question answering. *arXiv preprint arXiv:2004.04906* (2020).

[23] Jukka-Pekka Kauppi, Melih Kandemir, Veli-Matti Saarinen, Lotta Hirvenkari, Lauri Parkkonen, Arto Klami, Riitta Hari, and Samuel Kaski. 2015. Towards brain-activity-controlled information retrieval: Decoding image relevance from MEG signals. *NeuroImage* 112 (2015), 288–298.

[24] Omar Khattab and Matei Zaharia. 2020. Colbert: Efficient and effective passage search via contextualized late interaction over bert. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*. 39–48.

[25] Carol Collier Kuhlthau. 2005. Information search process. *Hong Kong, China* 7, 2005 (2005), 226.

[26] Kenton Lee, Ming-Wei Chang, and Kristina Toutanova. 2019. Latent retrieval for weakly supervised open domain question answering. *arXiv preprint arXiv:1906.00300* (2019).

[27] Jilong Li, Zhenxi Song, Jiaqi Wang, Min Zhang, and Zhiguo Zhang. 2024. BrainECHO: Semantic Brain Signal Decoding through Vector-Quantized Spectrogram Reconstruction for Whisper-Enhanced Text Generation. *arXiv preprint arXiv:2410.14971* (2024).

[28] Xiujun Li, Xi Yin, Chunyuan Li, Pengchuan Zhang, Xiaowei Hu, Lei Zhang, Lijuan Wang, Houdong Hu, Li Dong, Furu Wei, et al. 2020. Oscar: Object-semantics aligned pre-training for vision-language tasks. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXX 16*. Springer, 121–137.

[29] Nikos K Logothetis. 2008. What we can do and what we cannot do with fMRI. *Nature* 453, 7197 (2008), 869–878.

[30] Jiasen Lu, Dhruv Batra, Devi Parikh, and Stefan Lee. 2019. Vilbert: Pretraining task-agnostic visiolinguistic representations for vision-and-language tasks. *Advances in neural information processing systems* 32 (2019).

[31] Sean MacAvaney, Andrew Yates, Arman Cohan, and Nazli Goharian. 2019. CEDR: Contextualized embeddings for document ranking. In *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval*. 1101–1104.

[32] Niall McGuire and Yashar Moshfeghi. 2024. Prediction of the realisation of an information need: an EEG study. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2584–2588.

[33] Dominika Michalkova, Mario Parra-Rodriguez, and Yashar Moshfeghi. 2022. Information need awareness: an EEG study. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 610–621.

[34] Tom M Mitchell, Svetlana V Shinkareva, Andrew Carlson, Kai-Min Chang, Vicente L Malave, Robert A Mason, and Marcel Adam Just. 2008. Predicting human brain activity associated with the meanings of nouns. *science* 320, 5880 (2008), 1191–1195.

[35] Yashar Moshfeghi. 2021. Neurasearch: Neuroscience and information retrieval. In *CEUR Workshop Proceedings*, Vol. 2950. 193–194.

[36] Yashar Moshfeghi, Peter Triantafillou, and Frank Pollick. 2019. Towards predicting a realisation of an information need based on brain signals. In *The world wide web conference*. 1300–1309.

[37] Yashar Moshfeghi, Peter Triantafillou, and Frank E Pollick. 2016. Understanding information need: An fMRI study. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. 335–344.

[38] Rodrigo Nogueira and Kyunghyun Cho. 2019. Passage Re-ranking with BERT. *arXiv preprint arXiv:1901.04085* (2019).

[39] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. 2018. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748* (2018).

[40] Sakrapee Paisalnan, Frank Pollick, and Yashar Moshfeghi. 2022. Neural correlates of satisfaction of an information need. In *International Conference on Machine Learning, Optimization, and Data Science*. Springer, 443–457.

[41] Jerrin Thomas Panachakel and Angarai Ganesan Ramakrishnan. 2021. Decoding covert speech from EEG-a comprehensive review. *Frontiers in Neuroscience* 15 (2021), 642251.

[42] Ruiyang Ren, Yingqi Qu, Jing Liu, Wayne Xin Zhao, Qiaoqiao She, Hua Wu, Haifeng Wang, and Ji-Rong Wen. 2021. Rocketqav2: A joint training method for dense passage retrieval and passage re-ranking. *arXiv preprint arXiv:2110.07367* (2021).

[43] Motoshige Sato, Kenichi Tomeoka, Ilya Horiguchi, Kai Arulkumaran, Ryota Kanai, and Shuntaro Sasai. 2024. Scaling Law in Neural Data: Non-Invasive Speech Decoding with 175 Hours of EEG Data. *arXiv preprint arXiv:2407.07595* (2024).

[44] Robert S Taylor. 1968. Question-negotiation and information seeking in libraries. *College & research libraries* 29, 3 (1968), 178–194.

[45] A Vaswani. 2017. Attention is all you need. *Advances in Neural Information Processing Systems* (2017).

[46] François-Benoît Vialatte, Monique Maurice, Justin Dauwels, and Andrzej Cichocki. 2010. Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in neurobiology* 90, 4 (2010), 418–438.

[47] Zhenhailong Wang and Heng Ji. 2022. Open vocabulary electroencephalography-to-text decoding and zero-shot sentiment classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36. 5350–5358.

[48] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. 2002. Brain–computer interfaces for communication and control. *Clinical neurophysiology* 113, 6 (2002), 767–791.

[49] Lee Xiong, Chenyan Xiong, Ye Li, Kwok-Fung Tang, Jialin Liu, Paul Bennett, Junaid Ahmed, and Arnold Overwijk. 2020. Approximate nearest neighbor negative contrastive learning for dense text retrieval. *arXiv preprint arXiv:2007.00808* (2020).

[50] Ziyi Ye, Qingyao Ai, Yiqun Liu, Maarten de Rijke, Min Zhang, Christina Lioma, and Tuukka Ruotsalo. 2024. Generative Language Reconstruction from Brain Recordings. *arXiv preprint arXiv:2311.09889* (2024).

[51] Ziyi Ye, Qingyao Ai, Yiqun Liu, Maarten de Rijke, Min Zhang, Christina Lioma, and Tuukka Ruotsalo. 2024. Language Generation from Brain Recordings. (2024).