Cross-Sensory Comparison of EEG Signals for Brain-Based Information Retrieval

Niall Mcguire
niall.mcguire@strath.ac.uk
NeuraSearch Laboratory
University of Strathclyde
Glasgow, UK

Yashar Moshfeghi
yashar.moshfeghi@strath.ac.uk
NeuraSearch Laboratory
University of Strathclyde
Glasgow, UK

Abstract

Translating internal information needs into textual queries poses challenges for information retrieval, particularly for users with physical impairments or ill-defined search intentions. Brain Passage Retrieval (BPR) approaches map EEG signals directly to dense passage representations, bypassing text translation. Whilst recent work reports superior performance with auditory versus visual EEG, these findings emerge from unbalanced experimental conditions where sample sizes, vocabularies, and corpus characteristics differ substantially between modalities, making it unclear whether observed differences reflect genuine neural processing advantages or dataset artefacts. We investigate EEG-based retrieval performance comparing auditory and visual modalities under controlled dataset conditions. Using the Brennan (auditory, 49 subjects) and Nieuwland (visual, 51 subjects) datasets, balanced for vocabulary overlap and sample distributions, we train BPR models with transformer EEG encoders and BERT text encoders via contrastive learning. Under controlled conditions, we observe complementary performance characteristics: audio EEG demonstrates stronger recall (Hit@5: +74%, Hit@10: +140%) whilst visual EEG achieves better precision (Hit@1: +100%). These findings suggest that modality-specific strengths could inform the design of brain-computer interfaces for information retrieval.

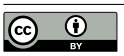
CCS Concepts

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

Keywords

Information Retrieval, EEG, Passage Retrieval, Brain-Computer Interfaces, Cross-Modal Learning

ACM Reference Format:
Niall Mcguire and Yashar Moshfeghi. 2026. Cross-Sensory Comparison of EEG Signals for Brain-Based Information Retrieval. In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '26), July 20–24, 2026, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3805712.3809922



This work is licensed under a Creative Commons Attribution 4.0 International License.
SIGIR '26, Melbourne, VIC, Australia
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2599-9/2026/07
https://doi.org/10.1145/3805712.3809922

1 Introduction

The translation of internal information needs into explicit textual queries represents one of the most enduring challenges in Information Retrieval (IR), where users must externalise often uncertain or ill-defined cognitive states through conventional interaction mechanisms [2, 22, 43]. This cognitive burden is particularly pronounced for users with physical impairments that limit traditional text input, where conventional interaction mechanisms may create substantial barriers to effective information access [60]. The complexity of this translation process often results in queries that incompletely or inaccurately represent users' true information needs, as users progress from visceral, unconscious needs to compromised forms suitable for IR systems [29, 52].

Recent advances in neuroscience have revealed promising approaches to addressing this fundamental gap through direct measurement of users' cognitive states during information interaction [16, 40, 42]. Electroencephalography (EEG) has emerged as a particularly viable neuroimaging modality for practical IR applications, offering high temporal resolution, portability, and significantly lower costs compared to fMRI or MEG systems [4]. Studies have demonstrated EEG's capability to detect core IR processes, from information need formation [35, 38] to relevance judgements [1, 50] and search satisfaction [47, 48]. Early EEG-based query formulation work treated brain decoding as machine translation [12, 57], but these approaches have been shown to rely on memorisation rather than meaningful brain-to-text mappings [23].

To address these limitations, Brain Passage Retrieval (BPR) has been proposed as an alternative framework that eliminates intermediate text translation by directly mapping EEG signals to dense passage representations in a shared semantic space [36, 37]. This approach leverages dense retrieval architectures adapted for neuroimaging data through specialised EEG encoders and cross-modal contrastive learning strategies [24, 26], achieving substantial improvements over existing EEG-to-text baselines whilst maintaining effectiveness across multiple participants.

Despite these advances in direct brain-to-passage mapping, a critical gap remains in our understanding of how different sensory modalities during information acquisition affect retrieval performance. The established literature on EEG-based IR has predominantly focused on visual stimulus presentations during reading tasks [19, 20], whilst the potential influence of auditory information processing on brain-based query representations remains underexplored. Cognitive neuroscience research suggests auditory and visual processing engage distinct neural pathways [18, 21], potentially affecting EEG-derived query representations.

Critically, previous investigations comparing EEG modalities have been confounded by substantial dataset imbalances. Studies have compared datasets with vastly different sample sizes (hundreds versus thousands of examples), divergent vocabulary distributions (minimal lexical overlap), distinct participant populations, and variable data collection protocols. These uncontrolled factors make it impossible to determine whether observed performance differences reflect genuine neural processing advantages of specific modalities or merely artefacts of dataset characteristics. No prior work has implemented rigorous balancing procedures to enable fair cross-modal comparison under controlled conditions.

In this work, we address this gap by conducting the first systematic comparison of retrieval performance between EEG signals acquired during auditory and visual stimulus presentation under balanced dataset conditions. Our key contribution is a comprehensive four-stage balancing procedure that controls for vocabulary overlap, word frequency distributions, and sample sizes, enabling fair comparison between modalities. We train separate BPR models on these balanced datasets and reveal complementary performance characteristics: visual EEG achieves superior precision whilst auditory EEG demonstrates stronger recall. These findings inform the design of adaptive brain-computer interfaces that leverage modality-specific strengths for information retrieval applications.

2 Related Work

The intersection of neuroscience and information retrieval has evolved from foundational studies of neural IR processes towards sophisticated brain-computer interfaces capable of direct semantic decoding. Early neuroimaging investigations established that core IR concepts manifest as distinct brain signatures detectable through various modalities [16, 43]. Moshfeghi et al. [41, 43] conducted seminal fMRI studies demonstrating that information need formation produces specific activation patterns in the posterior cingulate cortex. Michalkova et al. [38, 39] extended this work, characterising the metacognitive states preceding active search behaviour.

Building on this, EEG studies established consistent neurological correlates of relevance [1, 50, 51], and McGuire and Moshfeghi [35] achieved 90% accuracy predicting information need realisation before explicit query formulation.

These neurological insights have motivated the development of brain-machine interfaces for information retrieval. Early implementations utilised Steady-State Visually Evoked Potentials for character selection [7, 55], whilst more sophisticated approaches explored direct document recommendation [13, 14, 25]. Researchers have pursued direct EEG-to-text translation methods [12, 57, 61–63]. However, critical analyses by Jo et al. [23] identified fundamental limitations in these translation approaches, demonstrating that current EEG-to-text models rely heavily on teacher forcing [17] and exhibit similar performance with random noise inputs.

To address these translation limitations, the Brain Passage Retrieval framework proposed by McGuire and Moshfeghi [36, 37] represents a shift towards direct semantic alignment between brain signals and text passages. Critically, whilst McGuire and Moshfeghi [37] introduced auditory BPR and reported performance advantages over visual baselines, those comparisons were conducted under substantially uncontrolled conditions with divergent corpus sizes and vocabularies. The present work is the first to isolate genuine modality effects through rigorous balancing, directly addressing this methodological gap. Building upon dense retrieval architectures [24, 26], this method leverages contrastive learning principles [6, 46] to project EEG signals directly into dense passage representation spaces. The framework employs the Inverse Cloze Task [5, 30] to generate query-document pairs from naturalistic reading data.

However, existing BPR work has focused exclusively on visual stimulus presentation using the ZuCo datasets [19, 20], leaving the potential influence of auditory information processing unexplored. Cognitive neuroscience research has established that auditory and visual language processing engage distinct cortical pathways [10, 18], with auditory processing showing temporal dynamics and semantic organisation different from visual reading [21]. Critically, no prior work has implemented rigorous balancing procedures to enable fair cross-modal comparison under controlled conditions, making it impossible to distinguish genuine modality effects from dataset artefacts.

3 Methodology

3.1 Task Formulation

We investigate the comparative effectiveness of EEG signals captured during auditory and visual stimulus presentation within the Brain Passage Retrieval (BPR) framework [36, 37]. In BPR, a participant reads or listens to text whilst EEG signals are recorded; a model is then trained to map those neural signals directly to the correct passage in a corpus, treating the brain signal as a surrogate query without requiring the user to type anything. This eliminates intermediate text translation by directly mapping EEG signals to dense passage representations. During stimulus presentation (auditory listening or visual reading), EEG signals are recorded as $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_L] \in \mathbb{R}^{L \times D}$, where $\mathbf{x}_i \in \mathbb{R}^D$ represents the neural response to word w_i. Our framework employs dual encoders:

$$f_{\text{eeg}} : \mathbb{R}^{L \times D} \to \mathbb{R}^d \tag{1}$$

$$f_{\text{text}} : \mathcal{V}^* \to \mathbb{R}^d \tag{2}$$

where both map to a shared d-dimensional embedding space. Retrieval scores are computed as:

$$\text{sim}(\mathbf{q}, p) = f_{\text{eeg}}(\mathbf{q})^\top f_{\text{text}}(p) \tag{3}$$

Our central question is whether stimulus modality significantly affects EEG-derived query representations when datasets are rigorously balanced.

3.2 Datasets and Materials

We utilise two complementary datasets providing auditory and visual modality coverage. For auditory presentation, we employ the Brennan EEG dataset [3], containing recordings from 49 participants listening to the first chapter of *Alice's Adventures in Wonderland* (audio slowed 20%, normalised to 70dB SPL). For visual presentation, we use the Nieuwland dataset [45], with recordings from 51 participants reading discourse passages word-by-word (200ms display, 300ms blank). These datasets exhibit comparable properties: similar vocabulary sizes (Brennan: 2,129 tokens; Nieuwland:

2,247 tokens), balanced participant populations, and naturalistic comprehension paradigms. Both employed standardised EEG preprocessing: downsampling to 128 Hz, bandpass filtering (0.5–100 Hz), average reference re-referencing, and channel padding to 128 channels.

3.3 ICT Dataset Generation

We use the EEG-adapted Inverse Cloze Task framework[1] [5, 30] to generate query-document pairs. For each sentence, we extract query spans comprising 30% of sentence length with corresponding EEG signals, applying probabilistic masking ($p_{\text{mask}} = 0.9$) to the remaining document text. We generate a maximum of 2 ICT pairs per sentence with a minimum sentence length of 6 words and query length of 2–50 words, using fixed random seeds (seed=42) for reproducibility.

3.4 Dataset Balancing Procedure

Our four-stage balancing procedure controls for vocabulary overlap, word frequency distributions, and sample sizes—our primary contribution addressing confounding factors in previous cross-modal comparisons. **Stage 1: Vocabulary Filtering.** We apply dual filtering criteria to create comparable vocabularies whilst maintaining semantic diversity. First, query-level vocabulary filtering retains only ICT pairs where at least τ_{vocab} proportion of query words appear in the common vocabulary intersection V_{common}. This ensures queries are drawn from a shared lexical space. Second, word frequency ratio filtering removes words where:

$$\text{For } w \in V_{\text{common}} : \frac{\max(\phi_w^{\text{audio}}, \phi_w^{\text{visual}})}{\min(\phi_w^{\text{audio}}, \phi_w^{\text{visual}}) + \epsilon} \leq r_{\max} \tag{4}$$

where ϕ_w^{audio} and ϕ_w^{visual} denote the corpus frequency of word w in each modality, and $\epsilon = 10^{-6}$ prevents division by zero. Large frequency imbalances indicate different semantic usage patterns that could confound cross-modal comparison. **Stage 2: Stratified Sampling.** Following vocabulary filtering, we perform stratified sampling to create balanced datasets with equal sample sizes whilst preserving natural word frequency distributions. We categorise the filtered common vocabulary V_{filtered} into three frequency bins based on frequency rank (rank 1 = most frequent), providing coarse coverage of the frequency spectrum without over-parameterising the binning scheme:

$$V_{\text{high}} = \{w : \text{rank}(w) \leq |V_{\text{filtered}}|/3\} \tag{5}$$

$$V_{\text{medium}} = \{w : |V_{\text{filtered}}|/3 < \text{rank}(w) \leq 2|V_{\text{filtered}}|/3\} \tag{6}$$

$$V_{\text{low}} = \{w : \text{rank}(w) > 2|V_{\text{filtered}}|/3\} \tag{7}$$

For each bin b, we sample ICT pairs such that queries collectively contain words from frequency bin b in proportion $p_b = |V_b|/|V_{\text{filtered}}|$. This ensures both modalities exhibit matched frequency profiles across the spectrum [8]. **Stage 3: Final Equalisation.** We trim both datasets to equal sample sizes, controlling for the sample size imbalance that has plagued previous cross-modal EEG comparisons [20, 36, 37]. The balanced datasets are split into training, development, and test sets using stratified splitting with fixed random seeds (seed=42) to ensure reproducibility, following best practices in neural IR evaluation [53].

[1]Implementation taken from this GitHub repository

Table 1: Dataset characteristics before and after the four-stage balancing procedure. Balancing controls for corpus size, query count, vocabulary overlap, and word frequency distributions to enable fair cross-modal comparison.

Metric	Visual	Audio	Diff
Before Balancing			
Documents	615	592	23
Queries	814	814	0
Vocab overlap (%)	28		n/a
Freq similarity	0.32		n/a
After Balancing			
Documents	135	139	4
Queries	150	150	0
Vocab overlap (%)	**38**		+3 pp
Freq similarity	**0.51**		+0.19
Sample reduction	78%	77%	1 pp

3.5 Model Architecture

Our dual-encoder architecture employs frozen BERT-base-uncased [11] for text encoding, focusing adaptation on the EEG modality to avoid catastrophic forgetting [28]. The EEG encoder uses a 2-layer transformer with 8 attention heads and 768-dimensional hidden states. The encoder reshapes word-level EEG segments through linear projection:

$$\mathbf{H}_0 = \text{reshape}(\mathbf{X})\mathbf{W}_E + \mathbf{b}_E \tag{8}$$

where $\mathbf{W}_E \in \mathbb{R}^{(T\times C)\times d}$ and $\mathbf{b}_E \in \mathbb{R}^d$ project to hidden dimension $d = 768$. The sequence passes through transformer layers [54]:

$$\mathbf{H}_l = \text{TransformerLayer}_l(\mathbf{H}_{l-1}), \quad l \in [1, 2] \tag{9}$$

We investigate three pooling strategies (max, mean, CLS) for sequence-level aggregation [11, 24, 26], with final L2 normalisation.

Training employs InfoNCE contrastive loss [46]:

$$\mathcal{L}_{\text{contrastive}} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp(\text{sim}(\mathbf{q}_i, p_i)/\tau)}{\sum_{j=1}^{B} \exp(\text{sim}(\mathbf{q}_i, p_j)/\tau)} \tag{10}$$

where $\text{sim}(\mathbf{q}_i, p_j) = f_{\text{eeg}}(\mathbf{q}_i)^\top f_{\text{text}}(p_j)$ and $\tau = 0.07$ [56]. We use AdamW optimiser [33] (learning rate 10^{-4}, weight decay 0.01, batch size 32) with linear warmup (10 epochs) and decay, gradient clipping (norm 1.0) [49], and early stopping (patience 10). Models train up to 200 epochs on NVIDIA A100 GPUs. Separate models are trained per modality using identical hyperparameters.

4 Experimental Setup

We conduct a hyperparameter search to optimise model performance for each modality independently. Our search space includes pooling strategies (max, mean, CLS pooling) and training masking probabilities $p_{\text{mask}} \in \{0.25, 0.50, 0.75, 0.90\}$, which control the proportion of query words masked during document formation in the ICT framework. Additionally, we explore vocabulary overlap thresholds $\tau_{\text{vocab}} \in \{0.1, 0.3, 0.5, 0.7\}$ and word frequency ratio limits $r_{\max} \in \{20, 30, 50\}$ to assess balancing procedure sensitivity. All hyperparameter combinations are evaluated on the development set, with final models selected based on mean reciprocal rank (MRR).

Table 2: Cross-sensory EEG retrieval performance. Balanced conditions (bottom) use matched vocabulary distribution across modalities, enabling fair comparison. Unbalanced conditions (top, greyed) shown for reference only. Statistical significance from permutation test comparing balanced audio vs visual: $^{*}p < 0.05$, $^{**}p < 0.01$.

Modality	Cond.	#Docs	MRR	H@1	H@5	H@10
Visual	Unbal.	615	0.152	0.079	0.225	0.283
Audio	Unbal.	592	0.197	0.080	0.280	0.479
Visual	Bal.	135	0.244	**0.200****	0.280	0.320
Audio	Bal.	139	**0.281***	0.100	**0.487****	**0.767****
Audio vs Visual (Balanced)			↑15%	↓50%	↑74%	↑140%

Table 3: Pooling strategy ablation on balanced dataset (135–139 docs, 150 queries). Corpus sizes matched across modalities, enabling direct comparison. H@20 approaching 1.0 for audio [CLS] reflects ceiling behaviour expected at this corpus scale (top-20 covers ∼14% of 139 documents).

Configuration	MRR	H@1	H@10	H@20
Visual (Balanced, 135 docs)				
[CLS] pooling	0.164	0.113	0.233	0.273
Mean pooling	**0.244**	**0.200**	**0.320**	**0.373**
Max pooling	0.170	0.127	0.227	0.267
Audio (Balanced, 139 docs)				
[CLS] pooling	**0.281***	0.100	**0.767****	**0.993****
Mean pooling	0.183	0.087	0.393	0.491
Max pooling	0.239	**0.140**	0.440	0.550

We report performance on the held-out test set using standard retrieval metrics: MRR and Hit@k for k ∈ {1, 5, 10, 20}.

5 Results

Table 2 presents retrieval performance across balanced and unbalanced conditions. Under unbalanced conditions (greyed rows), audio EEG outperforms visual EEG (MRR: 0.197 vs 0.152), consistent with prior work [37]. However, substantial corpus size differences and divergent vocabularies confound direct comparison.

Under balanced conditions (non-greyed rows), where our four-stage procedure ensures matched corpus sizes (135 vs 139 documents), identical query counts (150 each), and controlled vocabularies, we observe complementary performance profiles. Audio EEG achieves superior overall effectiveness (MRR: 0.281 vs 0.244, +15%, $p < 0.05$) with substantially stronger recall: Hit@5 improves 74%, and Hit@10 improves 140% (both $p < 0.01$). Conversely, visual EEG demonstrates superior precision at top ranks (Hit@1: 0.200 vs 0.100, $p < 0.01$), suggesting audio EEG excels at broad retrieval coverage whilst visual EEG achieves greater top-rank accuracy. The superior recall of audio EEG may stem from extended stimulus presentation durations inherent to auditory processing, where longer acoustic exposure captures richer temporal dynamics compared to brief visual fixations [10, 18].

Table 3 reveals modality-specific optimal strategies: visual EEG performs best with mean pooling (MRR: 0.244) whilst audio EEG favours [CLS] pooling (MRR: 0.281, $p < 0.05$). This divergence suggests that semantic information distributes uniformly across visual representations but concentrates into classification tokens for sequential auditory processing [21, 54]. The consistency across pooling variants reinforces that observed differences reflect genuine cross-modal characteristics rather than hyperparameter sensitivity.

Table 1 demonstrates balancing effectiveness: near-perfect corpus matching (4-document difference), with vocabulary overlap improving from 28% to 38% and frequency similarity from 0.32 to 0.51. Whilst the 78% corpus reduction limits scale, this controlled comparison provides the first empirically valid cross-modal EEG retrieval assessment [36, 37].

6 Discussion and Conclusion

Our investigation faces limitations that suggest future directions. Dataset availability constrains analysis—only Brennan and Nieuwland provide sufficient overlap for fair comparison [3, 45], limiting generalisability. Balancing necessarily reduces corpus scale (135–139 documents, 150 queries), though sufficient for statistically significant results [9, 44]. Our English-only, Western-population focus suggests cross-lingual studies examining orthographic system effects [32] and diverse participant populations investigating neural variability [58]. Finally, the complementary profiles identified here motivate future investigation of multi-modal fusion strategies that combine visual precision with audio recall, which we leave as a direction for subsequent work [31, 34].

Future work should investigate neural mechanisms through frequency band analysis [15, 27], extend to other information access tasks, and conduct real-world deployment studies examining user acceptance and cognitive load [7, 60].

We present the first systematic investigation of EEG-based passage retrieval across auditory and visual modalities under controlled conditions. Through four-stage balancing, controlling vocabulary, frequency distributions, and corpus sizes, we reduce confounding factors found in previous comparisons. Our findings reveal complementary characteristics: audio EEG demonstrates superior recall (+74% Hit@5, +140% Hit@10) whilst visual EEG achieves better precision (+100% Hit@1), aligning with prior work [37] whilst confirming differences reflect genuine neural processing rather than dataset imbalances. These profiles have important implications: applications prioritising retrieval coverage should favour audio-based acquisition [59], whilst precision-critical applications benefit from visual acquisition. Hybrid systems could adaptively leverage complementary strengths. By establishing the first fair empirical comparison, our work provides foundational insights into cross-sensory EEG processing for information retrieval, informing the development of robust, accessible brain-computer interfaces.

Acknowledgments

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/W522260/1]. Results were obtained using the ARCHIE-WeSt High Performance Computer[2] based at the University of Strathclyde.

[2]www.archie-west.ac.uk

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] 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).

[6] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. 2020. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*. PMLR, 1597–1607.

[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] William Gemmell Cochran. 1977. *Sampling techniques*. john wiley & sons.

[9] Nick Craswell, Bhaskar Mitra, Emine Yilmaz, Daniel Campos, and Ellen M Voorhees. 2020. Overview of the TREC 2019 deep learning track. *arXiv preprint arXiv:2003.07820* (2020).

[10] 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.

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

[12] 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).

[13] 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.

[14] Manuel JA Eugster, Tuukka Ruotsalo, Michiel M Spapé, Ilkka Kosunen, Oswald Barral, Niklas Ravaja, Giulio Jacucci, and Samuel Kaski. 2014. Predicting term-relevance from brain signals. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. 425–434.

[15] 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.

[16] Jacek Gwizdka, Yashar Moshfeghi, and Max L Wilson. 2019. Introduction to the special issue on neuro? information science. *Journal of the Association for Information Science and Technology* 70, 9 (2019).

[17] Yongchang Hao, Yuxin Liu, and Lili Mou. 2022. Teacher forcing recovers reward functions for text generation. *Advances in Neural Information Processing Systems* 35 (2022), 12594–12607.

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

[19] 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.

[20] 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).

[21] 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.

[22] Peter Ingwersen. 1996. Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of documentation* 52, 1 (1996), 3–50.

[23] 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).

[24] 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).

[25] 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.

[26] 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.

[27] Kunjira Kingphai and Yashar Moshfeghi. 2023. On channel selection for EEG-based mental workload classification. In *International Conference on Machine Learning, Optimization, and Data Science*. Springer, 403–417.

[28] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, et al. 2017. Overcoming catastrophic forgetting in neural networks. *Proceedings of the national academy of sciences* 114, 13 (2017), 3521–3526.

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

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

[31] 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.

[32] Robert Litschko, Ivan Vulić, Simone Paolo Ponzetto, and Goran Glavaš. 2022. On cross-lingual retrieval with multilingual text encoders. *Information Retrieval Journal* 25, 2 (2022), 149–183.

[33] I Loshchilov. 2017. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101* (2017).

[34] 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).

[35] 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.

[36] Niall Mcguire and Yashar Moshfeghi. 2025. Towards brain passage retrieval: an investigation of EEG query representations. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 613–623.

[37] Niall McGuire and Yashar Moshfeghi. 2026. Cross-Sensory Brain Passage Retrieval: Scaling Beyond Visual to Audio. In *European Conference on Information Retrieval*. Springer, 481–496.

[38] 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.

[39] Dominika Michalkova, Mario Parra Rodriguez, and Yashar Moshfeghi. 2024. Understanding feeling-of-knowing in information search: An eeg study. *ACM Transactions on Information Systems* 42, 3 (2024), 1–30.

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

[41] Yashar Moshfeghi, Luisa R Pinto, Frank E Pollick, and Joemon M Jose. 2013. Understanding relevance: An fMRI study. In *Advances in Information Retrieval: 35th European Conference on IR Research, ECIR 2013, Moscow, Russia, March 24-27, 2013. Proceedings 35*. Springer, 14–25.

[42] 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.

[43] 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.

[44] Tri Nguyen, Mir Rosenberg, Xia Song, Jianfeng Gao, Saurabh Tiwary, Rangan Majumder, and Li Deng. 2016. Ms marco: A human-generated machine reading comprehension dataset. (2016).

[45] Mante S Nieuwland, Cas W Coopmans, and Rowan P Sommers. 2019. Distinguishing old from new referents during discourse comprehension: Evidence from ERPs and oscillations. *Frontiers in Human Neuroscience* 13 (2019), 398.

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

[47] Sakrapee Paisalnan, Frank Pollick, and Yashar Moshfeghi. 2021. Towards understanding neuroscience of realisation of information need in light of relevance and satisfaction judgement. In *International Conference on Machine Learning, Optimization, and Data Science*. Springer, 41–56.

[48] 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.

[49] R Pascanu. 2013. On the difficulty of training recurrent neural networks. *arXiv preprint arXiv:1211.5063* (2013).

[50] Zuzana Pinkosova, William J McGeown, and Yashar Moshfeghi. 2020. The cortical activity of graded relevance. In *Proceedings of the 43rd international acm sigir conference on research and development in information retrieval*. 299–308.

[51] Zuzana Pinkosova, William J McGeown, and Yashar Moshfeghi. 2022. Revisiting neurological aspects of relevance: An EEG study. In *International Conference on Machine Learning, Optimization, and Data Science*. Springer, 549–563.

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

[53] Nandan Thakur, Luiz Bonifacio, Maik Fröbe, Alexander Bondarenko, Ehsan Kamalloo, Martin Potthast, Matthias Hagen, and Jimmy Lin. 2024. Systematic evaluation of neural retrieval models on the Touché 2020 argument retrieval subset of BEIR. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1420–1430.

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

[55] 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.

[56] Tongzhou Wang and Phillip Isola. 2020. Understanding contrastive representation learning through alignment and uniformity on the hypersphere. In *International conference on machine learning*. PMLR, 9929–9939.

[57] 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.

[58] Chun-Shu Wei, Yuan-Pin Lin, Yu-Te Wang, Chin-Teng Lin, and Tzyy-Ping Jung. 2018. A subject-transfer framework for obviating inter-and intra-subject variability in EEG-based drowsiness detection. *NeuroImage* 174 (2018), 407–419.

[59] Ryen W White and Resa A Roth. 2009. *Exploratory search: Beyond the query-response paradigm*. Number 3. Morgan & Claypool Publishers.

[60] 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.

[61] 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).

[62] Ziyi Ye, Jingtao Zhan, Qingyao Ai, Yiqun Liu, Maarten de Rijke, Christina Lioma, and Tuukka Ruotsalo. 2024. Query Augmentation with Brain Signals. In *Proceedings of the 32nd ACM International Conference on Multimedia*. 7561–7570.

[63] Ruizhe Zhang, Qingyao Ai, Ziyi Ye, Yueyue Wu, Xiaohui Xie, and Yiqun Liu. 2024. Improving Legal Case Retrieval with Brain Signals. *arXiv preprint arXiv:2403.13242* (2024).