Supplementary Information for "High-level Prediction of Continuous Speech During Mind-Wandering"

Gal R. Chen[1], Rachel Finkelstein[2], Ariel Goldstein[3], Ran R. Hassin[1,4], Leon Y. Deouell[1,2]

1 Department of Psychology, The Hebrew University of Jerusalem

2 Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem

3 Department of Cognitive and Brain Sciences and Business School, The Hebrew University of Jerusalem

4 The Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem

*Corresponding author: Gal R. Chen

Email: galrefa.chen@mail.huji.ac.il

Table S1. Probability distribution of choices in mind-wandering probes

Option	Average Rate chosen
1 - Attentive to the chapter	0.720
2 - Daily matters	0.050
3 - My current state	0.102
4 - Personal concerns	0.019
5 - Daydreams	0.064
6 - External environment	0.029
7 - Other	0.016

Optional choices in mind-wandering probes (instructions were only read in Hebrew)

<u>Hebrew:</u>

1. הייתי קשוב/ה לפרק – **בכל מצב שבו העצירה הגיעה כשהיית מרוכז/ת בהאזנה, יש לבחור באופציה הזו.**
2. דברים יום-יומיים – מחשבות על על סדר היום שלך, משימות שאת צריכה עוד לבצע היום בלי קשר לניסוי
3. מצבי הנוכחי – המצב המנטלי או הפיזי שלך (חם לי, קר לי, האלקטרודה מגרדת לי וכו')
4. דאגות אישיות – מחשבות מדאיגות על העתיד, ובכללי על החיים
5. חלומות בהקיץ – פנטזיות או מחשבות לא תלויות מציאות
6. הסביבה החיצונית – דברים שקשורים בחלל שאת נמצאת בו כרגע, ואובייקטים בו: השולחן, הרמקולים הקירות וכו'
7. אחר – כל מחשבה אחרת שלא נכללת באופציות הנ"ל ואינה ריכוז בתוכן הפרק הנוכחי

<u>English</u>

1. I was attentive to the chapter - **Choose this option whenever the pause occurred while you were focused on listening.**
2. Daily matters - Thoughts about your daily schedule, tasks you need to complete today unrelated to the experiment
3. My current state - Your mental or physical condition (I'm hot, I'm cold, the electrode is itching)
4. Personal concerns - Worrying thoughts about the future, and generally about life
5. Daydreams - Fantasies or thoughts unrelated to reality
6. External environment - Things related to the space you're currently in and objects within it: the table, speakers, walls, etc.
7. Other - Any other thought not included in the above options and not related to concentrating on the current chapter's content

Table S2. Details on each audiobook used.

Audiobook	Duration (seconds)	Number of words	Number of probes	Mean surprise / negative log probability (SD)	Correlation of Mistral-based surprise with HeBERT (Pearson's r)
Urban legends	846.8	1476	19	2.75 (1.76)	0.679
Greek drama 1	1486.9	2816	28	2.74 (1.70)	0.574
Fiction writing	1249.5	2631	23	2.64 (1.72)	0.659
Greek drama 2	1168.8	2347	22	2.90 (1.71)	0.602
Product economy	1183.0	2284	22	2.59 (1.67)	0.670
Napoleon	1995.2	2088	22	2.57 (1.72)	0.583
Total	7930.2	13642	136	2.70 (1.72)	

Table S3. Results of within-subject mind-wandering classification based on PSD.

Band	Mean AUC	Mean Sensitivity	Mean Specificity	t-score (AUC vs 0.5)	df	p	corrected p
Delta (3-6 Hz)	0.53	0.651	0.946	1.287	24	.2103	1
Theta (6-8 Hz)	0.58	0.647	0.943	3.728	24	.001	.007
Alpha (8-12 Hz)	0.596	0.62	0.949	5.385	24	< .001	< .001
Low Beta (12-20 Hz)	0.591	0.676	0.942	4.257	24	< .001	.0021
High Beta (20-30 Hz)	0.547	0.611	0.942	1.992	24	.0579	.4053
All (3-30 Hz)	0.58	0.624	0.941	3.823	24	< .001	.0056

Note. p-values reported are two-sided and Bonferroni corrected.

Table S4.

Comparison of main analysis using 0.5-8 Hz and 0.5-20 Hz filters.

Analysis type	**MW/OT difference**	**0.5-8 Hz**		**0.5-20 Hz**		
		Cluster p-value	**Timepoints (s)**	**Cluster p-value**	**Timepoints (s)**	**Notes**
Word onset response	MW	< .001	0 – 0.95	< .001	0 - 0.96	
	OT	< .001	0 - 1	< .001	0 - 1	
	OT - MW	.007	0.05 - 0.37	.003	0.06 - 0.37	Direction: stronger early negativity for OT
Surprise response	MW	< .001	0.15 - 0.78	< .001	0.23 - 0.60	
	OT	< .001	0.24 - 0.76	< .001	0.24 - 0.67	
	OT - MW	0.486	0.16 - 0.28	.085	0.14 - 0.27	Direction: stronger early negativity for MW
Contextual Encoding	MW	< .001	-0.37 - 0.07	< .001	-0.39 - 0.05	1st cluster
	MW	< .001	0.13 - 0.49	< .001	0.11 - 0.37	2nd cluster
	OT	< .001	-0.63 - 0.83	< .001	-0.41 - 0.72	
	OT - MW	.023	0.03 - 0.15	.249	0.77 - 0.79	Difference not replicated

Note. MW: mind-wandering. OT: on-task. Reported are the clusters with maximum score from the cluster-based permutation analysis. The analysis is spatial-temporal for the rERP analyses, and temporal based on average of all channels for contextual encoding. p-values reported are one-sided and uncorrected across the 10 tests provided.

Table S5. Statistical Results for surprise response in different analyses

Time Window	Analysis Type	Mean OT	Mean MW	Mean Difference	Cohen's d	BF_{10}	t(24)	p
100-300ms	Only content words & high prior certainty words (z>-1)	-0.010	-0.087	0.078	0.441	0.978	1.892	0.071
	Control for word class and certainty	-0.017	-0.102	0.085	0.691	3.676	2.661	0.014
	Filter 0.5-20 Hz	-0.000	-0.081	0.081	0.705	2.932	2.540	0.018
300-500ms	Only content words & high prior certainty words (z>-1)	-0.125	-0.153	0.028	0.142	0.236	0.493	0.626
	Control for word class and certainty	-0.125	-0.149	0.023	0.191	0.264	0.703	0.489
	Filter 0.5-20 Hz	-0.122	-0.148	0.025	0.214	0.274	0.758	0.456

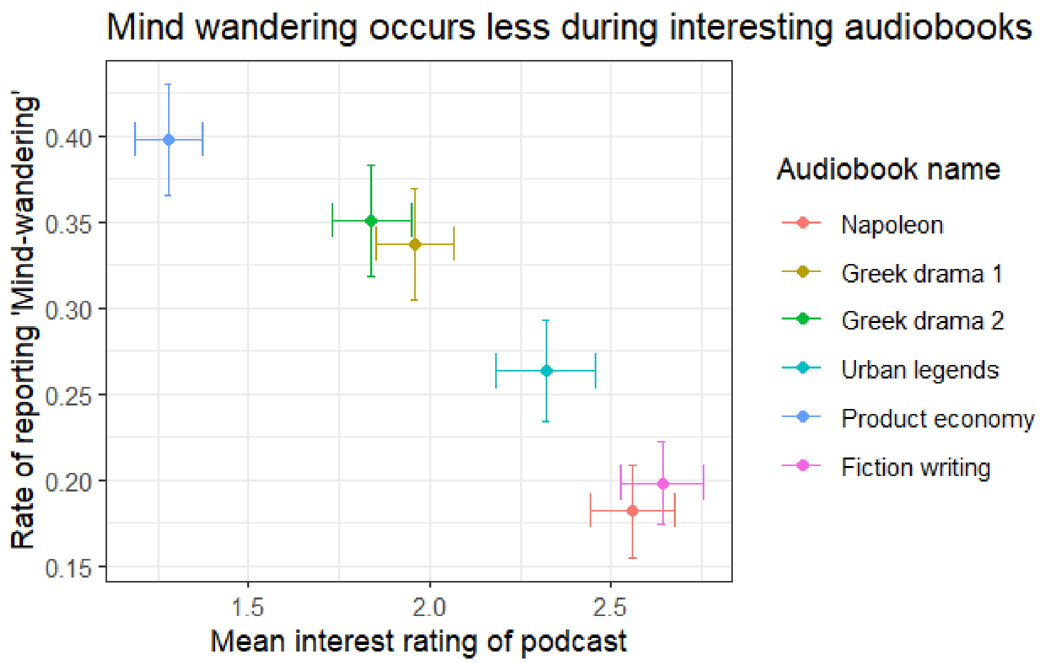
Note. MW: mind-wandering. OT: on-task. p-values reported are one-sided and uncorrected across the 6 tests provided here.

Table S6. Statistical Results for word onset response in different analyses

Time Window	Analysis Type	Mean OT	Mean MW	Mean Difference MW-OT	Cohen's d	BF_{10}	t (24)	p
100-300ms	Only content words & high prior certainty words (z>-1)	-0.123	-0.016	-0.108	-0.556	1.827	-2.276	0.032
	Control for word class and certainty	-0.179	-0.056	-0.123	-0.686	10.524	-3.188	0.004
	Filter 0.5-20 Hz	-0.180	-0.052	-0.128	-0.703	12.895	-3.285	0.003
300-500ms	Only content words & high prior certainty words (z>-1)	-0.150	-0.102	-0.048	-0.282	0.437	-1.281	0.212
	Control for word class and certainty	-0.239	-0.191	-0.049	-0.286	0.660	-1.617	0.119
	Filter 0.5-20 Hz	-0.242	-0.190	-0.053	-0.306	0.795	-1.751	0.093

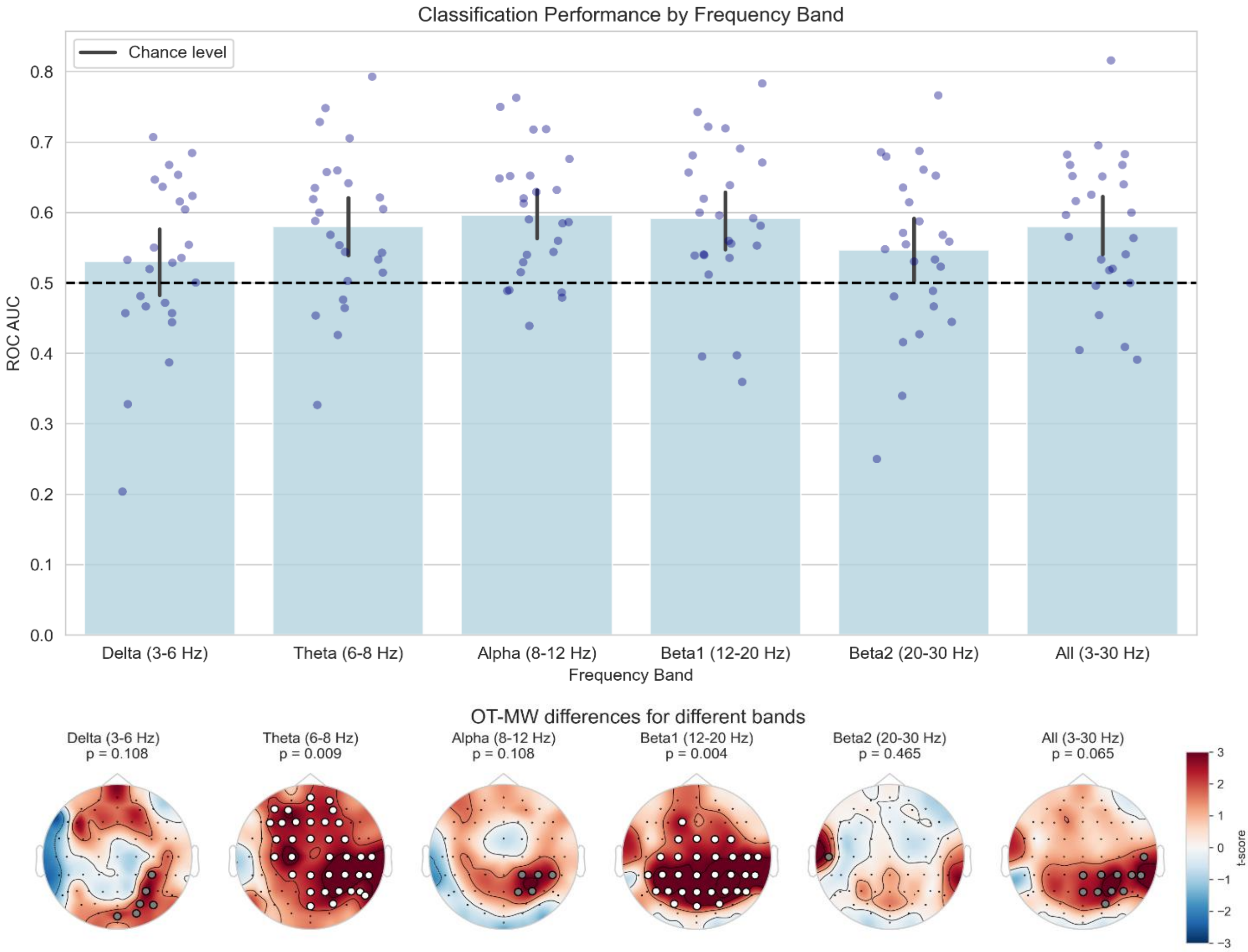
Note. MW: mind-wandering. OT: on-task. p-values reported are one-sided and uncorrected across the 6 tests provided.

Figure S1. Behavioral Analysis of mind-wandering reports



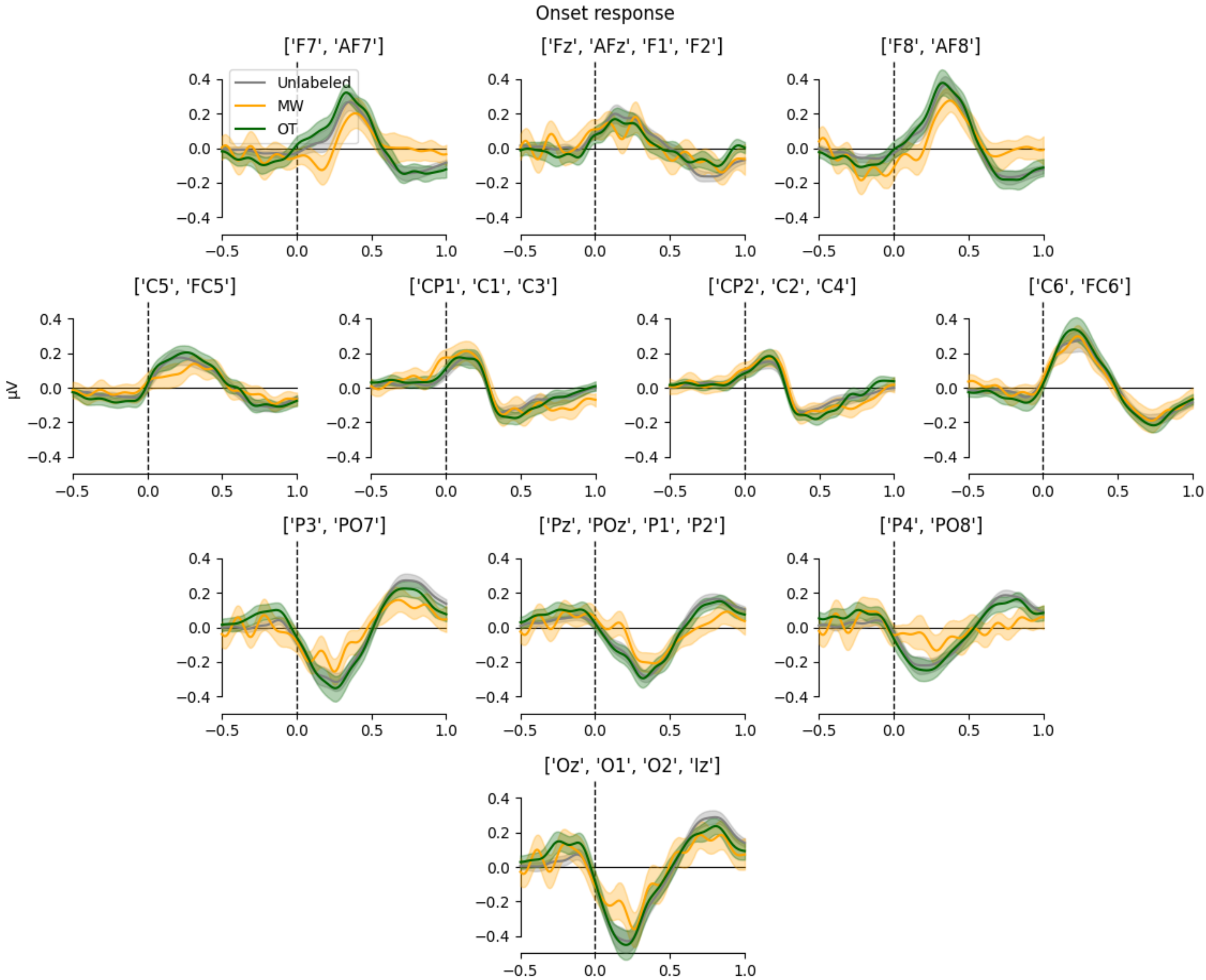
At the end of the Experiment, subjects were presented with the title of each chapter, alongside a short reminder of the topic, and were asked to rate how engaging the content of this audiobook chapter was. Options were 1 (Not interesting at all), 2 (Somewhat interesting), or 3 (Very interesting). We averaged the scores for each audiobook, and did the same for the rate of mind-wandering reported in each audiobook. The Pearson correlation between the two measures was r = 0.96, $t(4) = 7.2$, two-sided $p = 0.002$. Each point in the graph is one audiobook, and the error bars reflect standard error across subjects. The result indicates that more interesting audiobooks were also ones in which subjects were less likely to mind-wander.

Figure S2. Mind-wandering classification



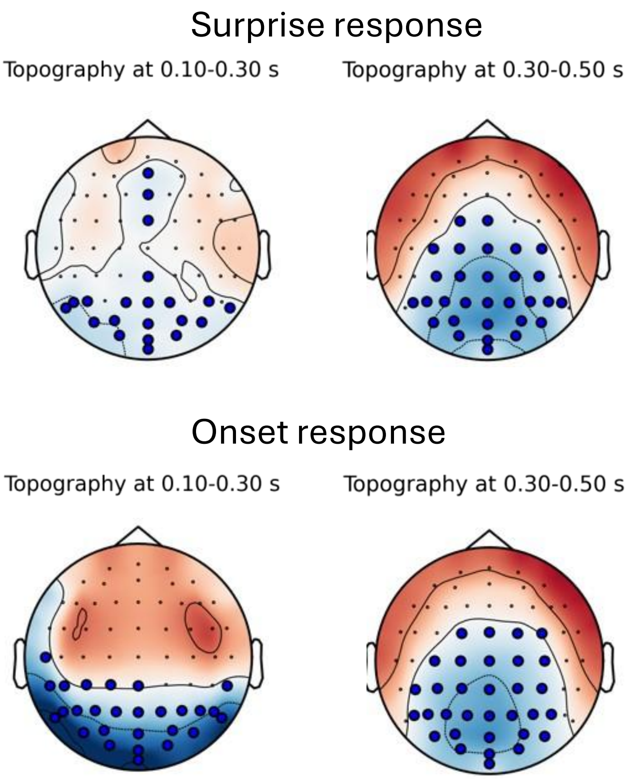
Top: Accuracy (ROC-AUC) of single-epoch (12 seconds) classification based on the average power in the epochs. The classification was based on a 10-fold cross-validation, with the average power in each band across all 64 channels. Error bars denote 95% confidence interval. Bottom panel shows the average difference in power between OT and MW reports, split between different bands. White circles denote channels belong to a significant (Bonferroni corrected one-tailed $p < .05$) cluster after spatial cluster-based testing. Grey circles show the maximum cluster (uncorrected one-tailed $p < .05$) for non-significant tests.

Figure S3. Responses to word onset in different clusters of spatially adjacent channels.



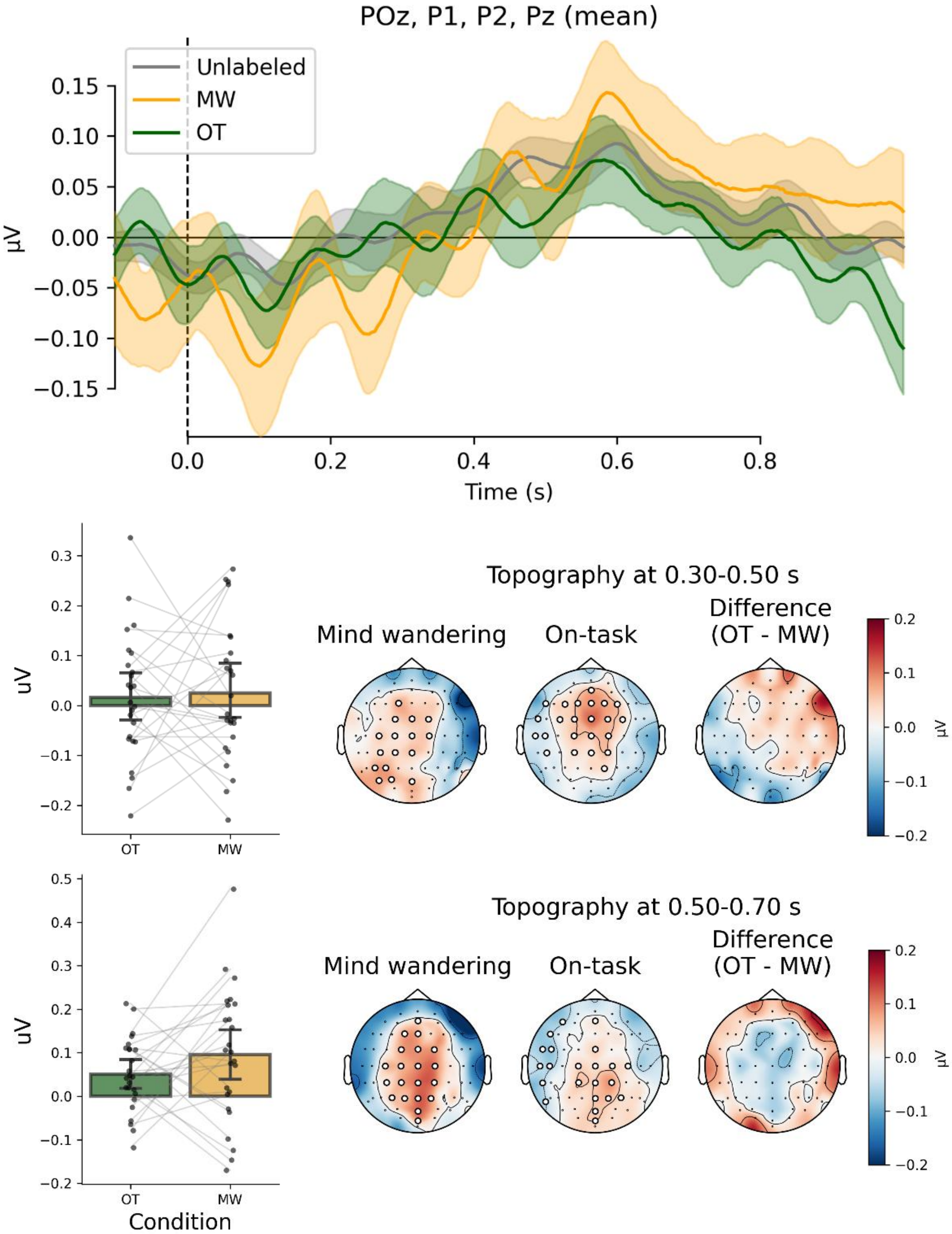
This figure is identical to Figure 3a, but each individual plot is based on a different cluster of channels.

Figure S4. Topography of surprise and onset responses for unlabeled data.



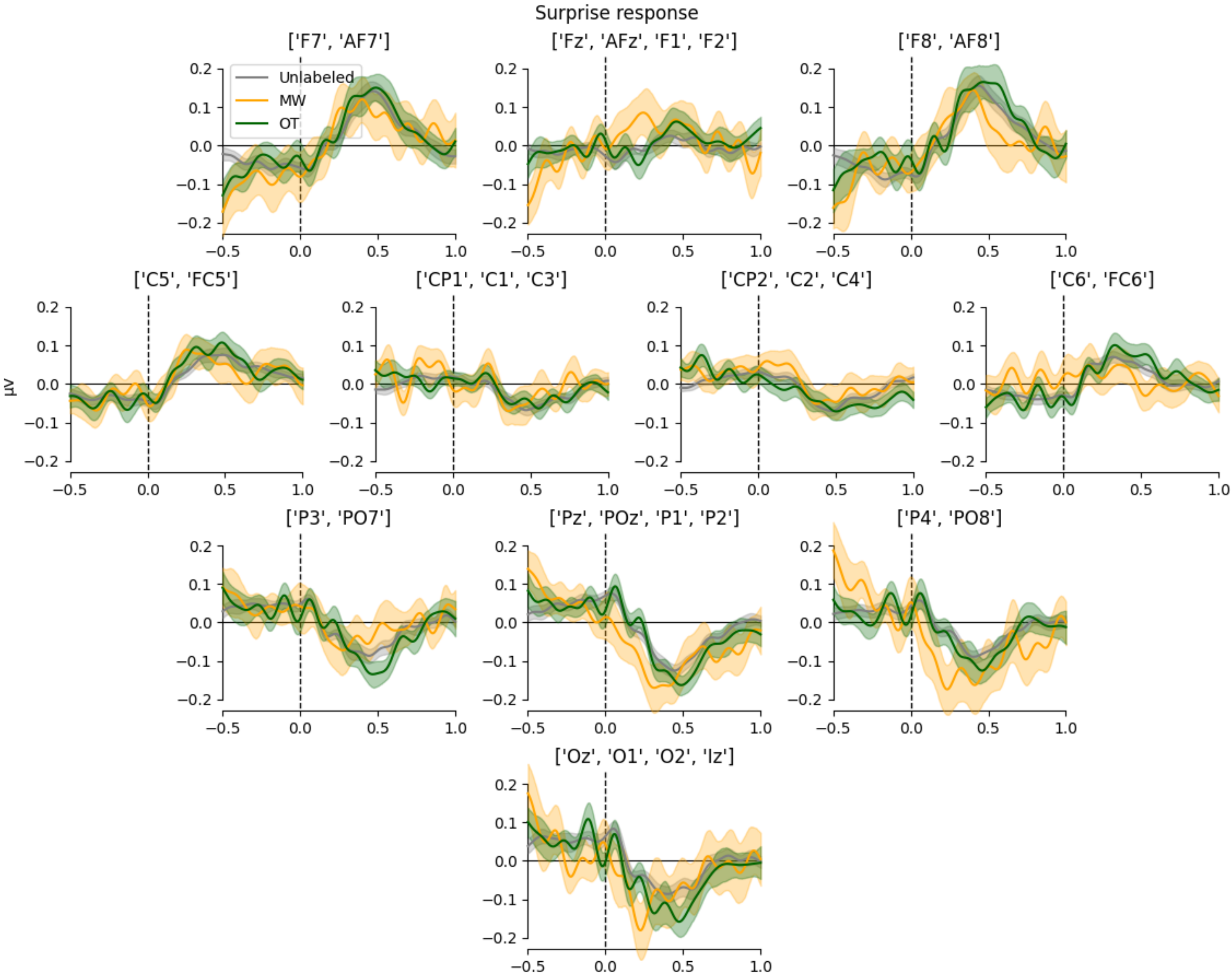
The marked dots are based on a similar spatio-temporal cluster permutation test conducted for mind-wandering and OT data in Figure 3. They are based on coefficients for which the design matrix zeroed for words belonging to mind-wandering/OT periods, and so they reflect another portion of the dataset.

Figure S5. Lexical frequency analysis



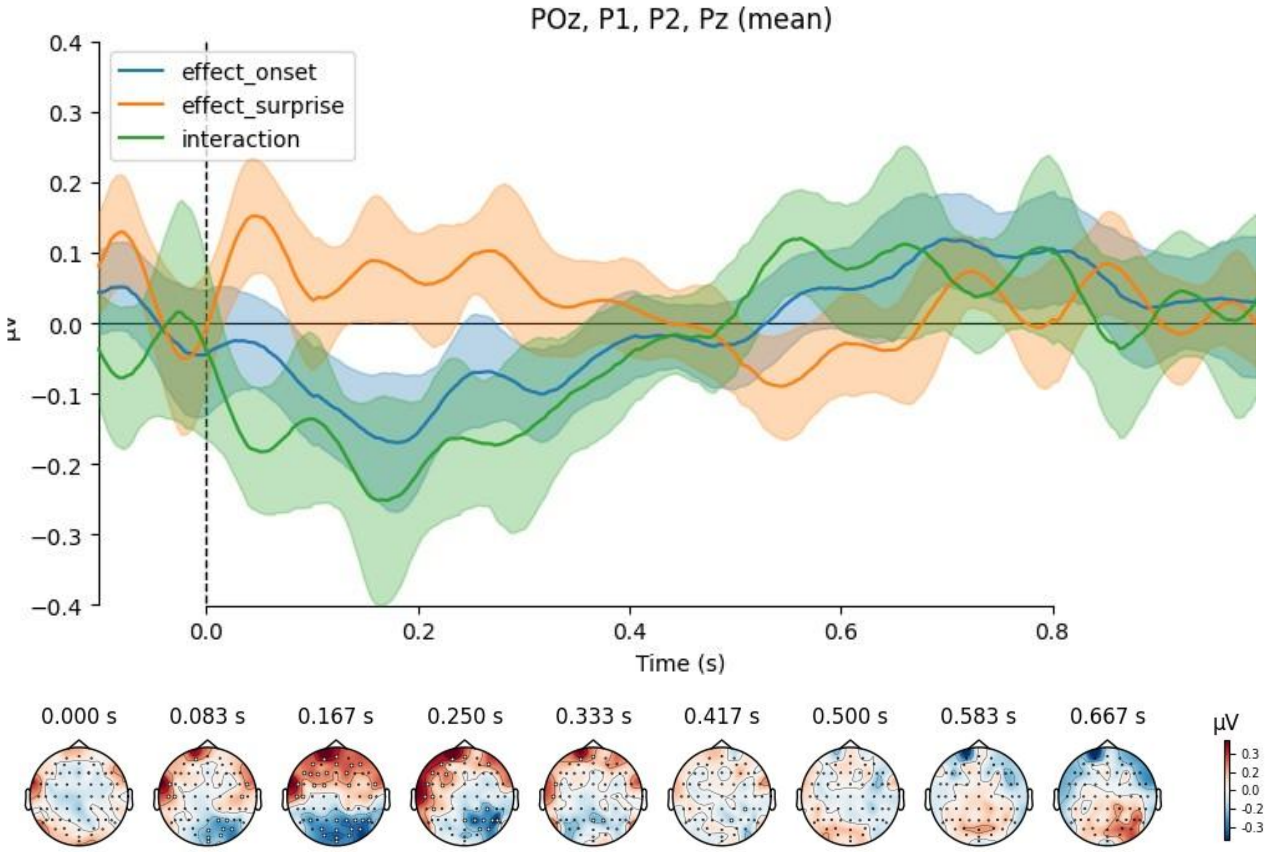
The figure presents the effect of lexical frequency, included in the model with surprise and other reported covariates, on the EEG response to words. It is identical to Figure 3 in its format. No significant differences were found in spatial-temporal cluster permutation test.

Figure S6. Responses to contextual surprise in different clusters of spatially adjacent channels.



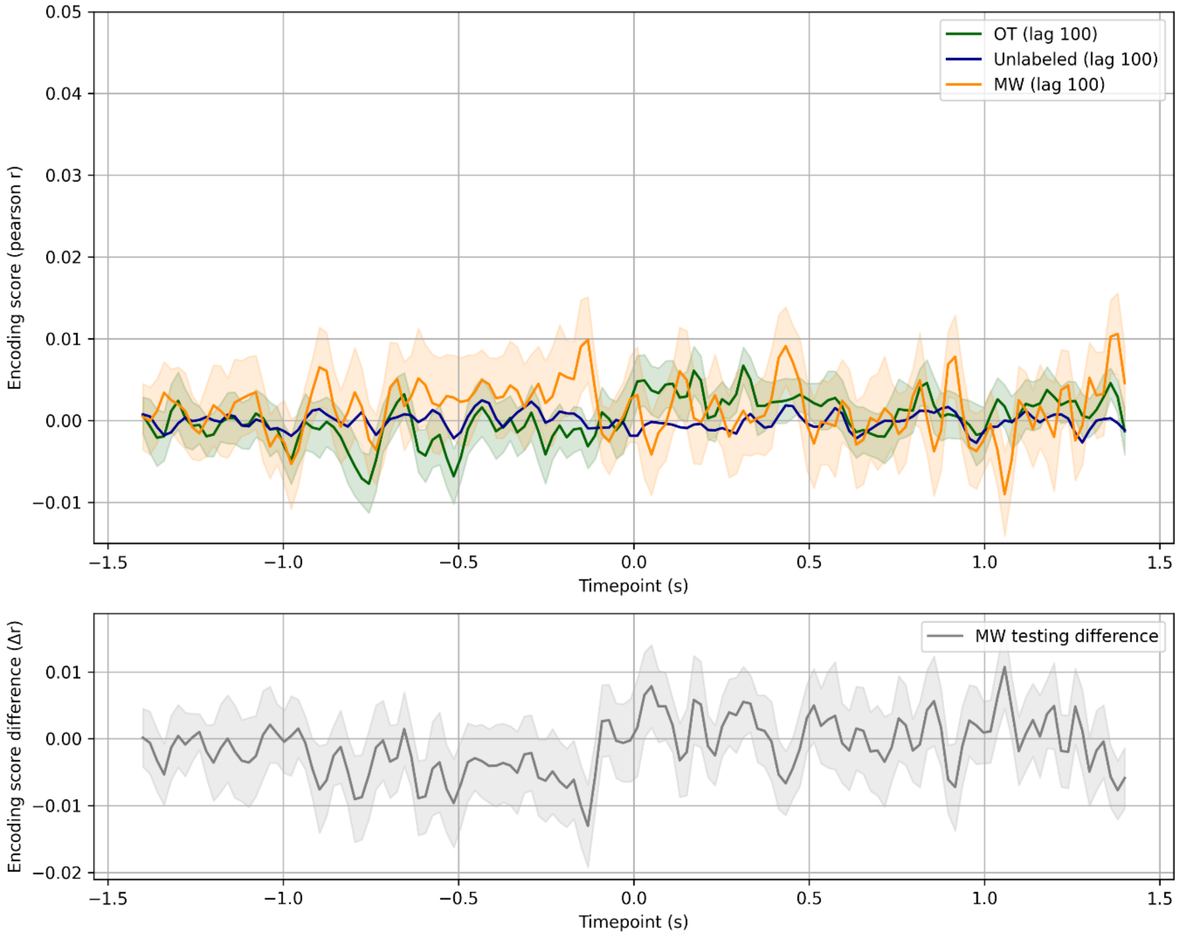
This figure is identical to Figure 3b, but each individual plot is based on a different cluster of channels.

Figure S7. The difference between mind-wandering effect on word onset and mind-wandering effect on contextual surprise.



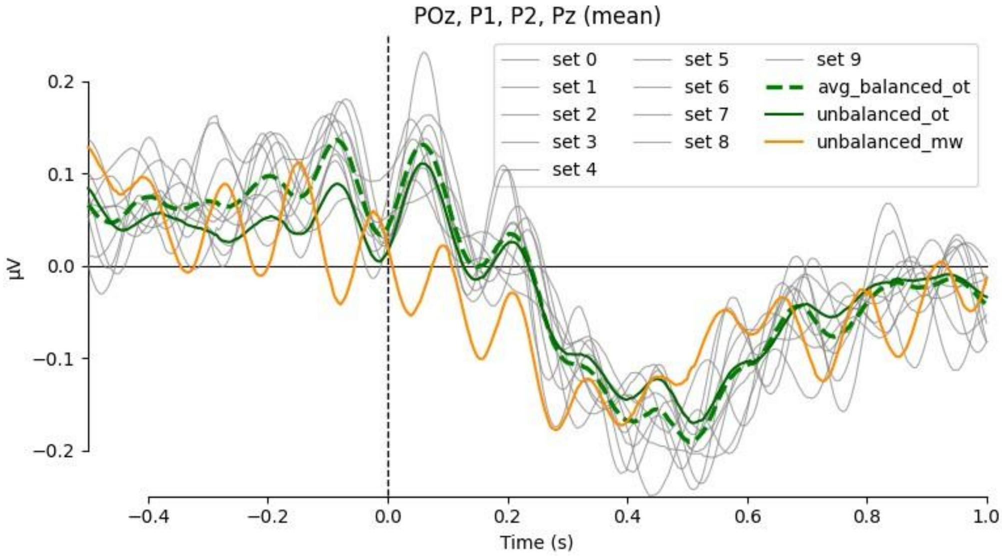
The top panel shows the average difference waves, based on the OT-MW difference for surprise (orange) and onset (blue) response. The green line shows the difference between surprise and onset effects (the interaction effect). The bottom panel shows the interaction difference wave topography, with the significant clusters in spatial-temporal cluster permutation test ($p < 0.05$) appearing in white dots.

Figure S8. Encoding of lagged contextual embeddings.



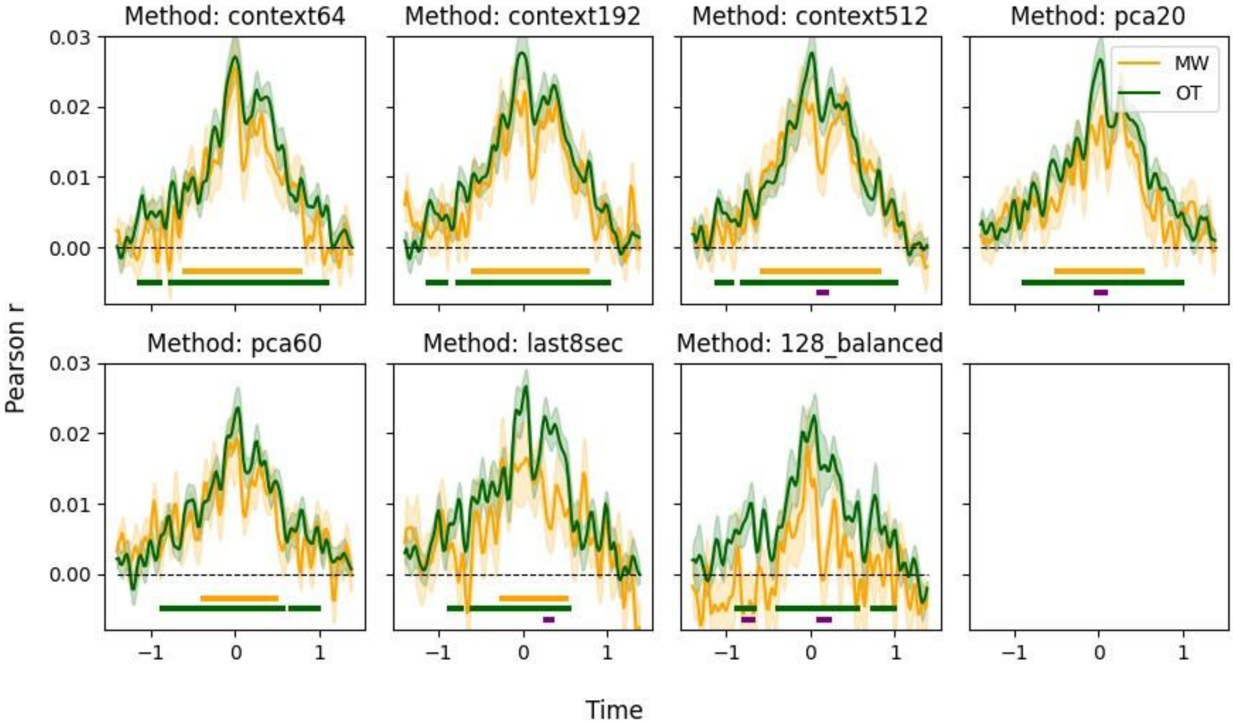
This analysis is identical to the one used in Figure 5 (contextual embedding analysis), but instead of using the embeddings of the word n for analyzing epoch n, we used the embedding of word n+100. As such, the last 100 word in each audiobook are discarded from this analysis.

Figure S9. Under-sampling of OT data in the regression analysis.



This analysis is similar to the one presented in Figure 3, but it is based on different sub-samples of the word in the OT category, so that the number of words is equivalent within each participant and response category (MW/OT). Events not in the sub-sample chosen in each of the 10 rounds were removed. Dashed green line is the average of all sub-samples. The solid lines illustrate the MW and OT effects presented in Figure 3 for reference (identical lines were observed for MW data here).

Figure S10. Alternative methods of extracting contextual embeddings.

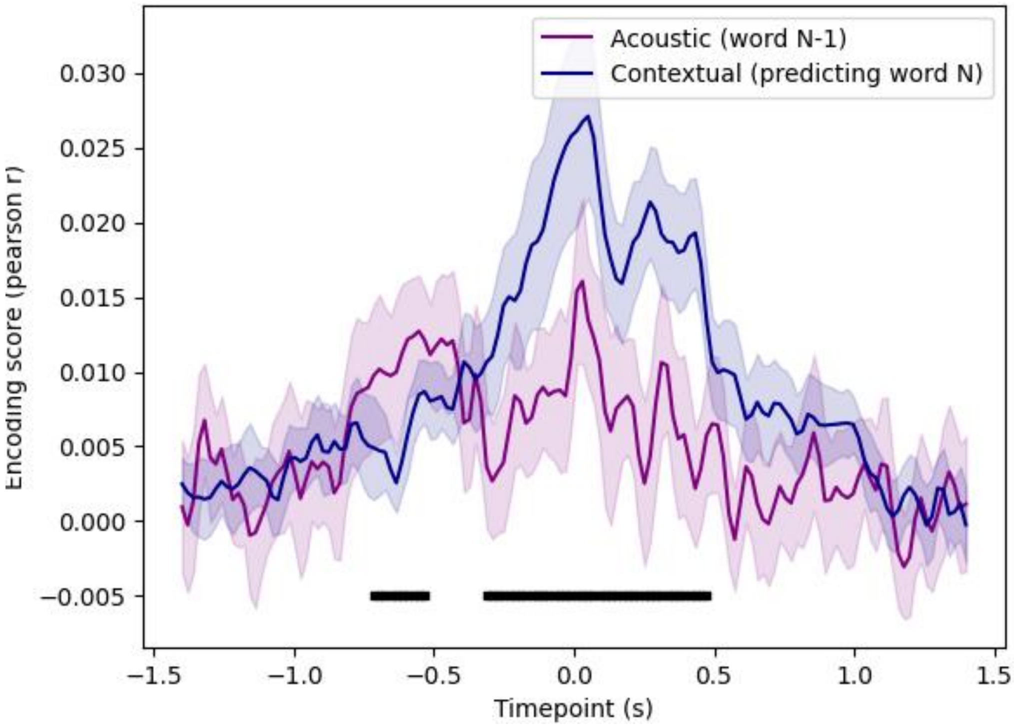


This figure is a reconstruction of Figure 5 with alternative preprocessing parameters. Solid horizontal lines reflect clusters with $p < .05$ for mind-wandering (MW, orange) and on-task (OT, green) states, and for the difference (purple). We tested a different number of components for reducing embeddings dimensionality (20/60 instead of 40), a different size of context window (64/192/512 instead of 128), and a different time window defined as MW (8 seconds instead of 12).

The balanced analysis includes oversampling of MW periods from out-of-test data, conducted to examine whether the training data was biased in favor of the more prevalent OT states. For a fully balanced training set based on strictly OT and MW periods, our effective training set will be diminished, on average, to ~20% its original size. To conduct a stronger analysis, we decided to supplement the original training set with out-of-test MW periods – oversampled several times to match the hypothesized rate of MW in the specific subject.

To illustrate the approach, if a participant had 9,000 unlabeled words, and reported MW in 20% of the time – this means approximately 1,800 of these words include MW periods, and we would need to oversample the (out of test) MW-labeled words up to 5,400 repetitions, for reaching a total of 7,200 MW-assumed words. This achieves the goal of providing a stable training set while reducing the presumed bias that would exist if representations were state-specific.

Figure S11. Encoding of acoustic embeddings.



This analysis is based on the first 40 principal components of mel-spectrogram at 0-350 ms post word onset, for all data without regards to subjective label. It shows that the acoustic information of the previous word cannot account for the time course of contextual encoding. Black squares denote timepoints in which the difference in encoding is significant at one-tailed $p < .05$ after cluster permutation testing. Both accuracies are based on 6-fold cross validation.