

Supplementary Information for

Reading comprehension in L1 and L2 readers: neurocomputational mechanisms revealed through
large language models

Supplementary Notes

Supplementary Note 1: Brain activation during reading for L1 and L2 readers

To complement our work, we conducted traditional univariate activation analysis to explore brain activation patterns during naturalistic reading. Following prior work^{1,2}, we established the individual-level fixed-effect model using the general linear model (GLM) in SPM12, with individual eye fixations as a regressor of interest. Regressors of no interest, including instructions and comprehension questions at the end of scanning, were added to ensure an unbiased implicit baseline^{3,4}. We also included head movement parameters and physiological regressors, and applied a high-pass filter with a cut-off of 100 seconds and the FAST algorithm⁵.

We established the group-level random-effect model and conducted one-sample T-tests to identify significant brain activations within each groups. The voxel wise threshold was set at $p < .001$, and $p < .05$ family-wise error (FWE) corrected at the cluster level. Subsequently, we conducted a two-sample T-test to compare brain activations between the groups. To determine the significance, we applied a peak-level family-wise error (FWE) threshold of $p < .05$ for multiple comparisons, with a minimum cluster size of 10 contiguous voxels. We used BrainNet Viewer for visualization⁶.

We identified significant brain activations in several regions for L1 readers including the left inferior occipital gyrus, left middle temporal gyrus, left precentral gyrus, right calcarine, right putamen, and bilateral medial frontal gyrus (see the left panel of Supplementary Figure 3a). Similarly, L2 readers demonstrated significant activations in the left middle occipital gyrus, left middle temporal gyrus, left precentral gyrus, left precuneus, right inferior occipital gyrus, and right calcarine (see the right panel of Supplementary Figure 3a). In line with previous research^{1,2,7}, L1 and L2 readers activated these brain regions supporting various linguistic

processes essential for naturalistic reading, such as single-word recognition and semantic integration.

Furthermore, L1 readers showed greater activations in bilateral cuneus and calcarine (see the left panel of Supplementary Figure 3b), brain regions often associated with lower-level orthographic processing and visual word identification^{1,2}, indicating more efficient visual word processing compared to L2 readers (see the right panel of Supplementary Figure 3b). In contrast, L2 readers displayed greater activations in the right inferior occipital gyrus, commonly involved in visuo-orthographic processing during reading⁸, and left precuneus, engaged in attentional control during visual processing^{7,9}, such as naturalistic reading. This heightened activation suggests an increased cognitive demand on visual analysis and the control of eye movements during naturalistic reading for L2 readers.

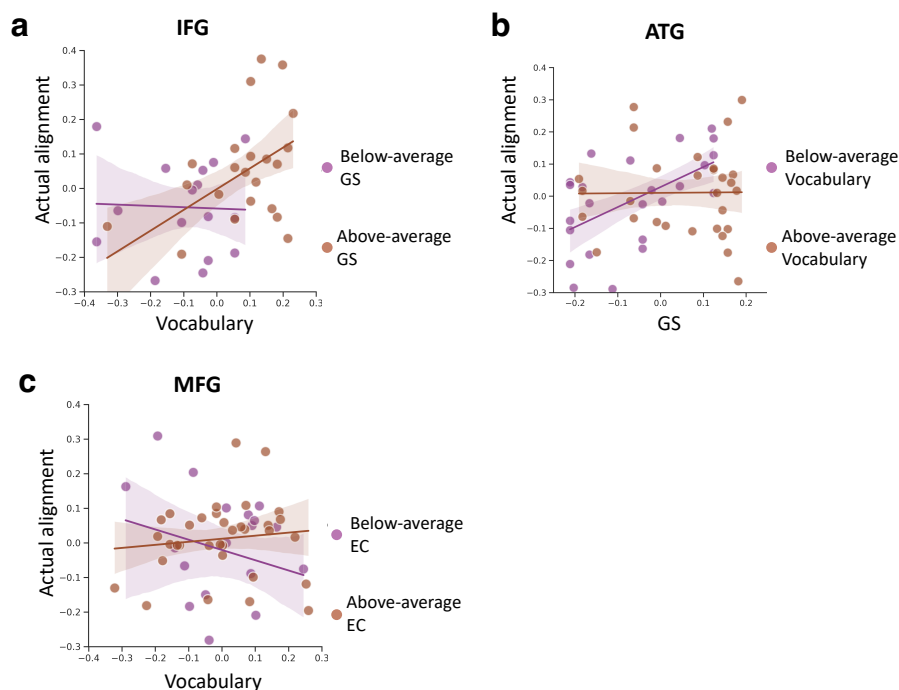
Supplementary Note 2: Model-brain alignment in language regions could not predict language background

To explore whether language background can be inferred from model-brain alignment, we built a classifier model to predict which population the reader comes from (i.e., L1 or L2 readers) using model-brain alignment in the language regions. Specifically, we constructed a classifier model using the Support Vector Machine (SVR) algorithm and leave-one-out cross-validation approach for individual language ROI. We found that neither of the model performance (ranging from 0.44 to 0.52; Supplementary Figure 2) of the classifier was significantly greater than the chance level (i.e., 50%) after FDR correction ($q < .05$). These results suggest that model-brain alignment in the language regions was unable to predict language background.

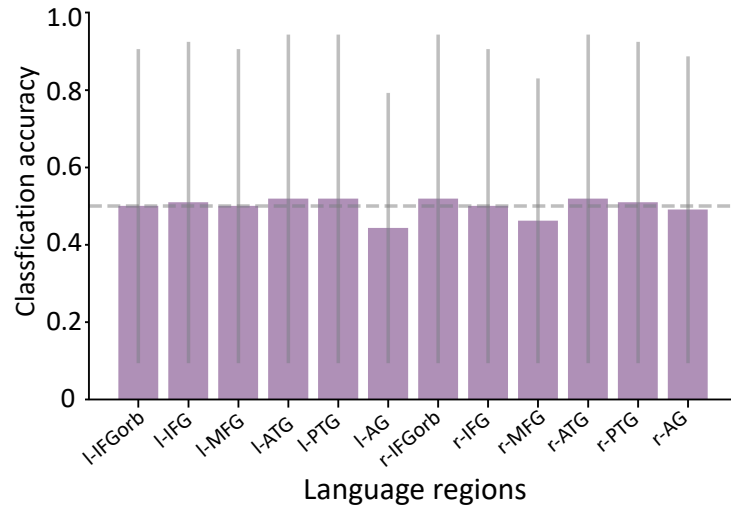
Supplementary Note 3: Comparable familiarity between L1 and L2 readers

We collected participants' familiarity with the text topics using a 5-point Likert scale. A two-sample t-test revealed that L1 ($M \pm SD: 2.67 \pm 0.69$; range: 1.5 ~ 4.20) and L2 readers ($M \pm SD: 2.70 \pm 0.90$; range: 1.0 ~ 4.8) showed comparable familiarity ($t(104) = -0.18, p = .85, d = -0.04$). In addition, to estimate the effect of familiarity on model-brain alignment, we established regression models including this measure as a predictor for each group, separately. We found that familiarity did not significantly predict the model-brain alignment ($p_s > .10$). These findings suggest that the two groups showed comparable familiarity with the text topics, which did not significantly impact the model-brain alignment.

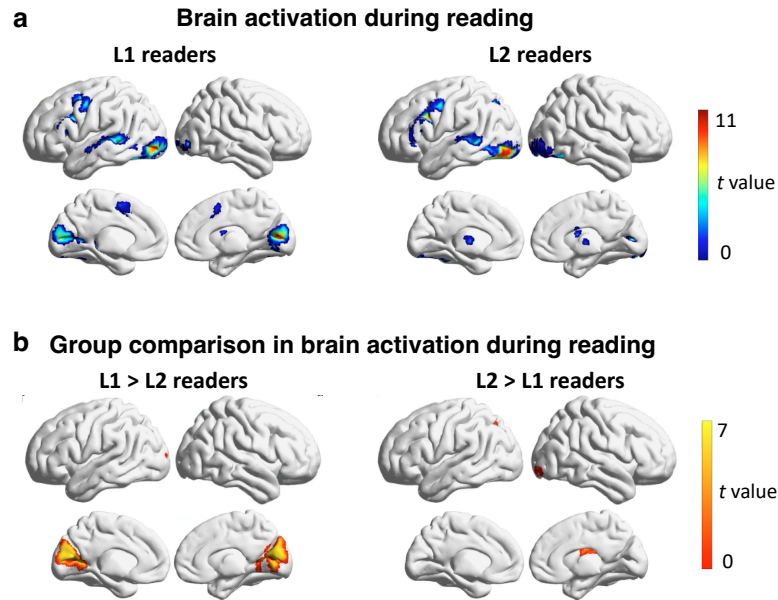
Supplementary Figures



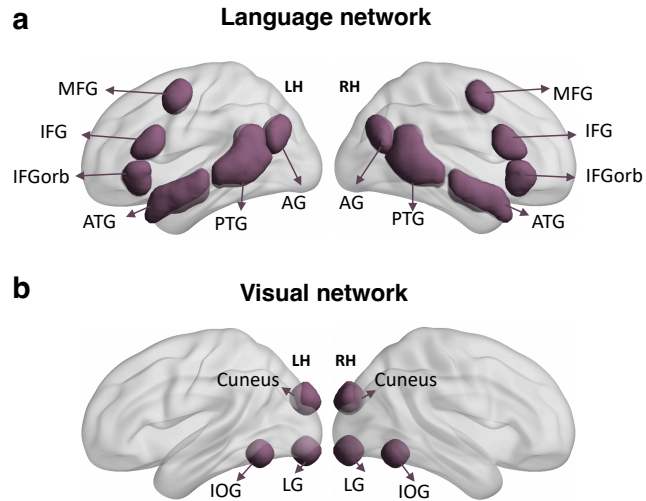
Supplementary Figure 1. Interaction effects among predictors. **a.** Interaction effects in the left IFG for L1 readers. The effect of general reading ability on model-brain alignment in the left IFG increased as vocabulary size increased. **b.** Interaction effects in the left ATG for L2 readers. The effect of vocabulary size on model-brain alignment in the left ATG decreased as general reading ability increased. **c.** Interaction effects in the left MFG for L2 readers. The effect of executive control ability on model-brain alignment in the left MFG decreased as vocabulary size decreased. *Note.* GR = General reading. EC = Executive control.



Supplementary Figure 2. The classification accuracy of model-brain alignment on language status. The gray line represents the 95% confidence level of the classification accuracy. Model-brain alignment could not predict language status (i.e., L1 or L2 readers) when the confidence level includes the chance level (i.e., 50%). **Note.** See a list of abbreviations in Supplementary Table 1.



Supplementary Figure 3. Univariate brain activation during naturalistic reading. a. Brain activation for L1 and L2 readers. **b.** Group comparison in brain activation between L1 and L2 readers. The left panel indicates greater brain activation for L1 readers than L2 readers, while the right panel represents greater brain activation for L2 readers than L1 readers.



Supplementary Figure 4. a. Language network. IFGorb: the orbital part of inferior frontal gyrus; IFG: inferior frontal gyrus; MFG: middle frontal gyrus; ATG: anterior temporal gyrus; PTG: posterior temporal gyrus; AG: angular gyrus. **b. Visual network.** IOG: inferior occipital gyrus; LG: lingual gyrus. LH: left hemisphere; RH: right hemisphere.

Supplementary Tables

Supplementary Table 1. Abbreviations of brain regions

Brain region	Abbreviation
angular gyrus	AG
anterior temporal gyrus	ATG
inferior frontal gyrus	IFG
orbital part of inferior frontal gyrus	IFGorb
inferior occipital gyrus	IOG
inferior parietal lobe	IPL
inferior parietal sulcus	IPS
lingual gyrus	LG
lateral prefrontal cortex	LPFC
middle frontal gyrus	MFG
middle prefrontal cortex	mPFC
orbital frontal cortex	OFC
posterior cingulate cortex	PCC
posterior temporal gyrus	PTG
superior parietal lobe	SPL
temporal gyrus	TG

Supplementary Table 2. Reading performance and behavioral assessments

Measures	L1 readers		L2 readers	
	M (SD)	Range (min ~ max)	M (SD)	Range (min ~ max)
Reading accuracy (%)	92.12 (5.48)	78.00 ~ 100.00	89.73 (8.64)	60.00 ~ 98.00
Total reading time (second/text)	84.71 (21.02)	42.39 ~ 163.59	109.00 (29.82)	57.62 ~ 170.14
Number of fixations	297.58 (56.22)	189.00 ~ 495.00	331.98 (64.47)	213.20 ~ 515.00
Mean fixation duration (second)	0.28 (0.03)	0.21 ~ 0.36	0.32 (0.05)	0.24 ~ 0.52
Vocabulary size	206.83 (10.05)	184.00 ~ 221.00	136.09 (37.44)	48.00 ~ 206.00
General reading ability	70.92 (23.34)	1.00 ~ 95.00	51.64 (33.37)	1.00 ~ 98.00
Altering network (millisecond)	44.90 (25.79)	7.00 ~ 140.00	51.25 (20.95)	10.00 ~ 98.00
EC network (millisecond)	108.90 (44.05)	36.00 ~ 300.00	94.20 (66.46)	31.00 ~ 534.00
Orienting network (millisecond)	38.76 (20.67)	3.00 ~ 95.00	37.64 (21.89)	-11.00 ~ 94.00

Note. EC = Executive control network.

Supplementary Table 3. STEM Texts Metrics

Text Metrics	Math	GPS	Mars	Circuit	Supertanker
# Words	308	307	310	304	302
# Sentences	28	28	31	30	31
# Avg. Words per Sentence	10.7	11.3	10.3	10.0	9.7

Supplementary Table 4. Ling model's prediction on model-brain alignment in L1 readers

Region	Independent variables	R^2	$Adj R^2$	F	β	t
Left IFGorb		.26	.20	4.06*		
	VS				0.35	2.19*
	GR				0.01	0.05
	VS * GR				2.39	2.48*
Left IFG		.32	.26	5.50*		
	VS				0.37	2.40*
	GR				0.08	.50
	VS * GR				2.62	2.85**
Left MFG		.28	.22	4.50*		
	VS				0.34	2.15*
	GR				0.12	.75
	VS * GR				2.28	2.40*
Left PTG		.25	.19	3.87*		
	VS				0.09	0.56
	GR				0.19	1.21
	VS * GR				2.94	3.03**

Note: VS and GR indicate vocabulary size and general reading ability, respectively. ** $p < .01$; * $p < .05$.

Supplementary Table 5. LingANT model's prediction on model-brain alignment in L2 readers

Region	Independent variables	R^2	$Adj R^2$	F	β	t
Left IFG		.34	.25	3.89*		
	VS				-.022	-1.48
	GR				0.37	2.52*
	Alerting				-0.41	-3.18**
	VS * GR				-1.71	-1.84
	VS * Alerting				-1.89	-1.49
	GR * Alerting				0.31	0.23
Left MFG		.31	.22	3.37*		
	VS				-.30	-1.93
	GR				0.48	3.21**
	Alerting				-0.29	-2.21*
	VS * GR				-1.55	-1.63
	VS * Alerting				-0.99	-0.76
	GR * Alerting				0.50	0.38
Left ATG		.29	.19	3.05*		
	VS				-0.08	-0.49
	GR				0.30	1.99 ⁺
	Alerting				-0.27	-2.07*
	VS * GR				-2.26	-2.34*
	VS * Alerting				-2.40	-1.83
	GR * Alerting				0.58	0.42
Left PTG		.27	.17	2.77 ⁺		
	VS				-0.23	-1.51
	GR				0.34	2.23*
	Alerting				-0.30	-2.19*
	VS * GR				-2.12	-2.17*
	VS * Alerting				-0.97	-0.73
	GR * Alerting				0.35	0.25
Left MFG		0.33	0.24	3.78*		
	VS				-0.36	-2.26
	GR				0.47	3.01**
	EC				-0.72	-2.17*
	VS * GR				-1.24	-1.27
	VS * EC				-3.97	-2.40*
	GR * EC				1.31	0.69

Note: VS, GR, and EC indicate vocabulary size, general reading ability, and executive control ability, respectively. $.050 < ^+p \leq .055$; $*p < .05$; $**p < .01$.

Supplementary Table 6. Full model's prediction on model-brain alignment in L2 readers

Region	Independent variables	R^2	$Adj R^2$	F	β	t
Left IFG		.44	.30	3.27*		
	LD				-.05	-0.33
	VS				-0.21	-1.34
	GR				0.37	2.50*
	Alerting				-0.42	-3.33**
	LD * VS				-2.83	-2.22*
	LD * GR				3.28	1.89
	LD * Alerting				2.79	1.79
	VS * GR				-1.29	-1.10
	VS * Alerting				-2.19	-1.44
	GR * Alerting				-1.12	-0.80
Left MFG		.40	.26	2.85*		
	LD				0.03	0.18
	VS				-0.29	-1.77
	GR				0.43	2.82**
	Alerting				-0.25	-1.97 ⁺
	LD * VS				-1.17	-0.90
	LD * GR				-0.17	-0.09
	LD * Alerting				4.14	2.58*
	VS * GR				-0.23	-0.19
	VS * Alerting				-3.00	-1.91
	GR * Alerting				0.06	-0.11
Left PTG		.39	.24	2.63 ⁺		
	LD				0.13	0.81
	VS				-0.25	-1.51
	GR				0.25	1.63
	Alerting				-0.24	1.84
	LD * VS				0.31	0.23
	LD * GR				-1.39	-0.77
	LD * Alerting				4.36	2.67*
	VS * GR				-0.97	-0.80
	VS * Alerting				-3.37	-2.12*
	GR * Alerting				0.38	0.26

Note: LD, VS, and GR indicate language dominance, vocabulary size, and general reading ability, respectively. ** $p < .01$; * $p \leq .050$; $0.50 \leq ^+p < .055$.

Supplementary References

1. Hsu, C.-T., Clariana, R., Schloss, B. & Li, P. Neurocognitive signatures of naturalistic reading of scientific texts: a fixation-related fMRI study. *Sci. Rep.* **9**, 1–16 (2019).
2. Schuster, S., Hawelka, S., Hutzler, F., Kronbichler, M. & Richlan, F. Words in context: The effects of length, frequency, and predictability on brain responses during natural reading. *Cereb. Cortex* **26**, 3889–3904 (2016).
3. Desai, R. H., Choi, W. & Henderson, J. M. Word frequency effects in naturalistic reading. *Lang. Cogn. Neurosci.* **35**, 583–594 (2020).
4. Schuster, S., Hawelka, S., Himmelstoss, N. A., Richlan, F. & Hutzler, F. The neural correlates of word position and lexical predictability during sentence reading: Evidence from fixation-related fMRI. *Lang. Cogn. Neurosci.* **35**, 613–624 (2020).
5. Todd, N. *et al.* Evaluation of 2D multiband EPI imaging for high-resolution, whole-brain, task-based fMRI studies at 3T: Sensitivity and slice leakage artifacts. *Neuroimage* **124**, 32–42 (2016).
6. Xia, M., Wang, J. & He, Y. BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One* **8**, e68910 (2013).
7. Zhou, W. & Shu, H. A meta-analysis of functional magnetic resonance imaging studies of eye movements and visual word reading. *Brain Behav.* **7**, e00683 (2017).
8. Cao, F., Young Kim, S., Liu, Y. & Liu, L. Similarities and differences in brain activation and functional connectivity in first and second language reading: Evidence from Chinese learners of English. *Neuropsychologia* **63**, 275–284 (2014).
9. Cauda, F. *et al.* Functional connectivity of the posteromedial cortex. *PLoS One* **5**, e13107 (2010).