Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Increased top-down semantic processing in natural speech linked to better reading in dyslexia

Anastasia Klimovich-Gray [a,*], Giovanni Di Liberto [b], Lucia Amoruso [c,e], Ander Barrena [d], Eneko Agirre [d], Nicola Molinaro [c,e]

[a] *University of Aberdeen, Aberdeen, UK, AB24 3UB*
[b] *University of Dublin, Trinity College, Dublin, Ireland, D02 PN40*
[c] *BCBL, Basque Center on Cognition, Brain and Language, San Sebastian, Spain, 20009*
[d] *HiTZ - Ixa, University of the Basque Country UPV/EHU, San Sebastian, Spain, 48940*
[e] *Ikerbasque, Basque Foundation for Science, Bilbao, Spain, 48009*

ARTICLE INFO

Keywords:
Dyslexia
Magnetoencephalography
Speech
Semantics
Predictive processing

ABSTRACT

Early research proposed that individuals with developmental dyslexia use contextual information to facilitate lexical access and compensate for phonological deficits. Yet at present there is no corroborating neuro-cognitive evidence. We explored this with a novel combination of magnetoencephalography (MEG), neural encoding and grey matter volume analyses. We analysed MEG data from 41 adult native Spanish speakers (14 with dyslexic symptoms) who passively listened to naturalistic sentences. We used multivariate Temporal Response Function analysis to capture online cortical tracking of both auditory (speech envelope) and contextual information. To compute contextual information tracking we used word-level Semantic Surprisal derived using a Transformer Neural Network language model. We related online information tracking to participants' reading scores and grey matter volumes within the reading-linked cortical network. We found that right hemisphere envelope tracking was related to better phonological decoding (pseudoword reading) for both groups, with dyslexic readers performing worse overall at this task. Consistently, grey matter volume in the superior temporal and bilateral inferior frontal areas increased with better envelope tracking abilities. Critically, for dyslexic readers only, stronger Semantic Surprisal tracking in the right hemisphere was related to better word reading. These findings further support the notion of a speech envelope tracking deficit in dyslexia and provide novel evidence for top-down semantic compensatory mechanisms.

1. Introduction

Cortical speech processing is dynamic. Brain circuits adapt their processing to background noise levels (Hervais-Adelman et al., 2012; Khalighinejad et al., 2017), attentional demands (Ahveninen et al., 2011) and speaker characteristics (Sjerps et al., 2019). This adaptation is marked by real-time changes in cortical patterns and is arguably supported by predictive coding mechanisms enabling the listener to flexibly up or down weight top-down (contextual) versus bottom-up (perceptual) cues (Hakonen et al., 2017; Donhauser and Baillet, 2019). In addition to online adaptation, there is evidence for longer-term changes in cortical speech processing due to ageing (Broderick et al., 2021) or learning new skills such as musical training or another language (Tervaniemi et al., 2022). Crucially, this functional plasticity of the cortical circuits that process speech could be key for preserved speech comprehension in developmental disorders of language.

Developmental dyslexia is associated with a phonological deficit linked to poor perceptual tracking of the speech signal (Goswami, 2011; Lizarazu et al., 2018; Molinaro et al., 2016). Despite a wealth of evidence that phonological skills play a major role in spoken language processing, dyslexic readers typically do not display severe speech comprehension impairments (Nation and Snowling, 1998; Shaywitz et al., 1999). Early behavioural work suggested that semantic and contextual information can partially compensate for lower-level deficits (Stanovich, 1980) by facilitating lexical and semantic levels of word recognition in sentence contexts. There is substantial behavioural evidence that poor and dyslexic readers rely on contextual facilitation during spoken (Chiappe et al., 2004) and written (Corkett and Parrila, 2008; Frith and Snowling, 1983) language comprehension. Critically, others have shown greater contextual facilitation for adult dyslexic compared to typical readers (Nation and Snowling, 1998) with a similar pattern of results also shown for children with dyslexia (van der Kleij et al. 2019).

Neuroimaging allows us to simultaneously explore the spatiotemporal cortical dynamics of both lower-level auditory and higher-level contextual semantic processes. Neurophysiological studies show marked differences between typical and dyslexic readers in auditory and phonological processing but comparable neural markers of contextual semantic facilitation (N400 component related to contextual uncertainty), both for spoken and written words (Bonte and Blomert, 2004; Helenius et al., 1999; Jednoróg et al., 2010; Mody et al., 2008; Rüsseler et al., 2007; Silva-Pereyra et al., 2003). Very few studies show reduced effects of context (diminished N400 effect) for dyslexic word reading (Schulz et al., 2008). This evidence is comparable with the behavioural literature and suggests robust contextual semantic processing in dyslexia, despite persistently poor phonological skills. Although these neuroimaging findings are compatible with compensatory semantic mechanisms in dyslexia, the evidence linking related neural markers to specific compensatory mechanisms is missing.

Here we used a novel combination of multivariate encoding analyses of the MEG data and state-of-the-art computational linguistics analysis to explore the cortical tracking of lower-level auditory and higher-level contextual information for natural speech in typical versus dyslexic readers. We related these processes to reading abilities in both groups. Finally, we related these measures with the structural morphology of the reading network. Auditory information processing was measured by the cortical tracking of the speech envelope. To estimate participants' ability to track contextual semantics we derived contextual Semantic Surprisal for words in sentences. Surprisal has previously shown robust neural correlates and was linked to the N400 response for less predictable words (Gillis et al., 2021; Broderick et al., 2021). To estimate Semantic Surprisal we used word-level probabilities produced with the Transformer Neural Network, which is currently the most accurate machine learning tool for this task. Envelope and Surprisal measures were used for the multivariate encoding analysis on time-resolved MEG cortical responses to capture dynamic neural signatures of related processes.

In the present study in addition to the main analysis of the MEG responses we also explored how variability in cortical speech tracking is related to structural properties of the cortical network linked to reading. Learning to read facilitates development of dedicated cortical circuits encompassing bilateral temporal, temporoparietal, occipitotemporal and inferior frontal areas. In typically developing readers, increases in grey matter volume (GMV) within this network are related to better reading skills (Hoeft et al., 2007). Accordingly, dyslexic individuals show reduced GMV in occipitotemporal areas of the ventral visual stream (left fusiform which includes visual word form area) supporting letter and text processing and in the temporoparietal areas related to audio-visual information integration (Krafnick et al., 2014). Apart from atypicalities in the visual ventral stream, developmental dyslexia has been related to a lack of left-hemispheric bias (i.e., left>right) in the superior temporal grey matter volume (Dole et al., 2013), which could explain the reduced ability to track auditory information (Lizarazu et al., 2015). However, in a recent meta-analysis, Ramus et al. (2018); see also Peyre et al. (2020) highlighted the difficulty in differentiating dyslexic from normal readers based on neuroanatomical estimates. The aim of the exploratory GMV analysis is to provide further complementary evidence for the relation between speech processing and reading.

In this study we focused on adults, who, unlike children, have had more time and experience to develop semantics-based compensatory strategies. Following others (Stanovich, 1980; Chiappe et al., 2004) we reasoned that if dyslexic readers use top-down compensatory strategies, they should be present in speech comprehension, where the adverse effects of phonological deficits initially emerge. If this strategy is also beneficial for reading, top-down contextual effects (Semantic Surprisal tracking) in speech processing should be correlated with reading performance and, potentially, variability in the GMV. We therefore predict that reading skills will be positively correlated with the Semantic Surprisal tracking, especially in the dyslexic reader group. Following behavioural findings (Nation and Snowling, 1998; van der Kleij et al.,

2019), it is also possible that dyslexic readers will show overall greater Surprisal tracking than controls despite weaker Envelop tracking, implying a speech processing strategy that more heavily relies on contextual analysis during word recognition.

2. Methods

2.1. Participants

We collected data from 43 right-handed native Spanish speakers (14 male) with no known history of neurological disorders or hearing issues and normal or corrected-to-normal vision. Data for 2 participants was discarded due to poor signal quality. Average age was 30.3 years, SD = 9.78. This experiment was approved by the Ethics Committee of the Basque Centre on Cognition, Brain and Language (BCBL) and all participants signed a written informed consent form and were paid for their time. To verify the presence of dyslexic symptoms in those who reported having reading issues or previous dyslexia diagnosis, participants were administered a battery of tests which included: WAIS IQ test, PROLEC-SE-R (Spanish reading test for adolescents), RAN (rapid automatic naming) and several additional tests of phonological abilities (phonological deletion, phonological short-term memory). From these tests, 14 participants (mean age 30.57, SD = 10.05, 4 male) were confirmed as presenting dyslexic symptoms (see section below for details). This sample size of the dyslexic group is comparable with multiple previously published neuroimaging papers (Molinaro et al., 2016, - 10 dyslexic adults; Power et al., 2016 - 12 dyslexic children; Hämäläinen et al., 2012 - 11 dyslexic adults) and was previously shown to be sufficient to demonstrate replicable differences between control and dyslexic groups in synchronisation to the speech envelope. We therefore considered this sample size sufficient for the planned analyses.

2.2. Behavioural tests of dyslexic symptoms

Our main group included both dyslexic and control participants. During participant recruitment we invited both those with official dyslexia diagnosis (diagnosed as children or adolescents) and those who suspected they had dyslexia due to poor reading skills but were undiagnosed. To confirm dyslexic symptoms, we administered the Spanish reading test (PROLEC-SE-R), which is a comprehensive assessment of various reading-related skills (Cuetos et al., 2014). As our experiment involved adults, we reasoned that although some reading related skills (such as semantic categorisation, grammatical fluency and narrative comprehension) may be within the normal range due to extended reading experience, basic word and pseudoword reading would still be affected by the underlying dyslexic symptoms. Therefore, we calculated three separate measures from the PROLEC-SE-R and used scores on any one of these measures to determine dyslexic symptoms. These measures were: (1) the overall PROLEC-SE-R test score, (2) Word reading and (3) Pseudoword reading scores. To assess if each dyslexic individual significantly deviated from the normative control sample ($n = 20$), reading efficiency scores were computed for Words and Pseudowords on the basis of both reading time and accuracy as (accuracy/time)*100 and then converted to t values, adjusted for a small sample comparison (Crawford and Howell 1998). Our dyslexic sample inclusion criteria were that an individual's t value was less or equal to the critical t value of -1.72 (alpha level 0.05, df 19, one-tailed) for at least one of the three key measures. Specifically, we ensured that all participants that were placed into the dyslexia group showed either (a) very poor phonological skills at -1.72 score for Pseudo-word reading efficiency (alpha level 0.05, df 19, one-tailed) or (b) showed overall very poor reading skills at score -1.72 or below for Word reading efficiency or PROLEC-SE total score (alpha level 0.05, df 19, one-tailed) accompanied by a mild phonological deficit at -1.32 or less for Pseudoword reading (alpha level 0.1, df 19, one-tailed). Our criteria were designed to make sure that the dyslexia group included people with mild to severe phonological deficits. Pseudoword reading is

a robust test of phonological abilities in dyslexic adults who may have partially compensated for word-level reading deficits. With these criteria our dyslexia dataset contained 14 individuals. We confirmed that all dyslexic participants had a score of 80 points or above on the intelligence quotient (WAIS) test and on each part of the non-verbal IQ index (Perceptual Organization and Processing Speed). Finally, to ensure that our normative control sample of $n = 20$ participants had word and pseudoword reading skills that approximated the population level we compared their data with previously published data (Molinaro et al., 2016) for a normative sample of $n = 46$ native Spanish readers who completed the same Word and Pseudoword reading tasks. In this larger control sample the Word efficiency score was $M = 187.77$ SD $= 28.02$ and the Pseudoword $M = 114.63$ SD $= 13.54$. These numbers are not significantly different from the smaller normative sample in this study ($t = 0.14$, $p = 0.88$ and $t = 0.39$, $p = 0.69$, respectively). Efficiency scores for Word ($M = 117.15$, SD $= 24.3$) and Pseudoword reading ($M = 66.64$, SD $= 15.09$) in our dyslexia sample are significantly lower compared to this larger normative sample ($t = 8.54$, $p < 0.0001$).

Poor reading, however, can also be related to generally poor language comprehension skills. In the latter case we do not expect that poor readers will use contextual comprehension as a strategy (Nation and Snowling, 1998). We confirmed that none of the 14 dyslexic participants in our sample had significant narrative comprehension deficits (adjusted $t > -1.72$, $p < 0.05$) compared to the normative control sample. For planned group-level comparisons (t-test) of the neural activity in the dyslexic participants with that of typical readers we selected 14 controls matched with the dyslexic participants for age and non-verbal IQ (measured by a WAIS Matrix test). For the regression analysis where neural activity was correlated with reading-related scores the whole sample ($n = 41$) was used. The key behavioural measures for the full group ($n = 41$) as well as the matched Control and Dyslexia groups are presented in Table 1 below: Total WAIS IQ test score (maximum score 155), Matrix component of the WAIS test (maximum score 26), PROLEC-SE-R comprehensive reading skills test with Word and Pseudoword reading sub-components separately.

2.3. Experimental design

2.3.1. Stimuli

Participants listened to a total of 160 audio sentences (recorded on a Marantz PMD digital recorder) spoken by a native female Spanish speaker at normal conversational pace (average sentence length – 13.2 words, SD = 3.2). All sentences were selected from a collection of on-line Spanish text corpora to include a varied degree of semantic constraints (News Corpora - 63 M (Bojar et al. 2014); esCow - 150 M sentences (Schäfer and Bildhauer, 2012); Billion Word Corpus - 1.5B words (Cardellino, 2016) and Wikicorpus - 120 M words (Padró et al., 2010)). This large collection of corpora was needed to train the neural network and later derive word-level Semantic Surprisal measures (next section). During data acquisition participants listened to an additional 160 sentences that were modified to change their original speed and a list of 80 unconnected words (the current study was part of a larger project on speech processing in typical and atypical populations – see Klimovich-Gray et al., 2021 for details) - both were discarded from the current analysis.

The 160 unmodified sentences had a varied degree of constraint with respect to target nouns embedded in them (target noun position in sentence $M = 9.3$, SD = 3, for full details of stimuli design see Klimovich-Gray et al., 2021), resulting in 80 more and 80 less constraining sentences. The degree of constraint was measured as the Surprisal (negative log2 of the probability - Hale, 2016) associated with the target word (all nouns) embedded in the sentence, given preceding words. If Surprisal is large the given word is not strongly expected from the preceding content (e.g. target word 'kitchen' in the sentence: 'This studio also has its own separate kitchen'). If Surprisal is relatively low, the context strongly constrains a given target word (e.g. target word 'kitchen' in the sentence: 'Breakfast was served in the kitchen every morning.'). We ensured that there was a reasonable range of semantic constraint across sentences - average Surprisal of a given target was 14.4 (SD = 2.34) in the weakly constraining sentences and 2.5 (SD = 1.33) in the strongly constraining sentences. Finally, all sentences were manually checked by native Spanish speakers who ensured that no sentences contained strong emotional or graphic content.

2.3.2. Experimental procedure

Each participant performed 3 blocks (fixed order) of data acquisition. In the first block resting state MEG activity was recorded as participants looked at the blank screen for 5 min with their eyes open (resting state data not analysed here). The second block, which was the focus of this study, consisted of the experimental sentences which participants were instructed to listen to attentively. To ensure attention during this block, participants occasionally (25% of trials) answered a simple yes/no comprehension question about the immediately preceding sentence, while looking at the black fixation cross. Participants answered with an index finger button press and the hand (right vs left) used for yes/no response was counterbalanced across participants. The third block consisted of passive listening to 80 single words (which were not used in the current analysis). All auditory stimuli were delivered with a random inter-stimulus interval (ISI) (from 1 to 2.5 s) via non-magnetic plastic tubes.

2.3.3. Data acquisition

The MEG data was acquired in a magnetically shielded room with a whole-scalp system (Elekta Neuromag, Helsinki, Finland) and the band-pass filter set to 0.01 – 330 Hz, 1 kHz sampling rate. Subjects' head positions were continuously monitored with four Head Position Indicator (HPI) coils. Coil position was digitised relative to the anatomical fiducials (nasion, left and right preauricular points) with a 3D digitizer (Fastrak Polhemus, Colchester, VA, USA). Subjects' horizontal and vertical eye movements and heart rate were monitored using bipolar electrodes. A subset of participants (neurotypical readers, $n = 20$, mean age = 28.78, SD = 8.89; dyslexic readers, $n = 13$, mean age = 30.07, SD = 10.18) also underwent structural MRI scanning in a single session, using a 3.0 Tesla Siemens Magnetom Trio Tim scanner (Siemens AG, Erlangen, Germany), located at the BCBL in Donostia-San Sebastián. A high-resolution T1-weighted scan was acquired with a 3D ultra-fast gradient echo (MPRAGE) pulse sequence using a 32-channel head coil and

Table 1
Behavioural assessment for Control and Dyslexic participants.

	Main group ($N = 41$)	Matched Controls ($N = 14$)	Dyslexia group ($N = 14$)
Age	$M = 30.02$, SD = 9.73	$M = 29.28$, SD = 8.54	$M = 30.57$, SD = 10.05
IQ Total	**$M = 103$, SD = 13.35**	**$M = 109.8$, DS=11.3**	**$M = 96$, SD = 11.8**
IQ Matrix test	$M = 20.09$, SD = 3.85	$M = 21.76$, DS=2.02	$M = 20.07$, SD = 3.62
PROLEC-SE-R	**$M = 96.02$, SD = 18.4**	**$M = 108.57$, SD = 11.71**	**$M = 76.64$, SD = 10.42**
Word reading	**$M = 162.35$, SD = 46.9**	**$M = 185.92$, SD = 37.8**	**$M = 117.15$, SD = 24.3**
Pseudoword reading	**$M = 95.67$, SD = 29.3**	**$M = 113.07$, SD = 22.45**	**$M = 66.64$, SD = 15.09**

Variables showing a significant difference between Controls and Dyslexic readers are in bold, $p < 0.05$ two-tailed.

with the following acquisition parameters: FOV = 256; 160 contiguous axial slices; voxel resolution $1 \times 1 \times 1$ mm^3; TR = 2300 ms, TE = 2.97 ms, flip angle = 9°.

2.3.4. MEG data processing and analysis

Data pre-processing and epoching was done using the open source MNE Python platform and analysis pipeline (Gramfort et al., 2014). First, we used MaxFilter 2.2 to perform signal-noise separation. Temporal extension of the signal space separation (Taulu et al., 2005) was applied to separate external noise from head-internal signal. Noisy and flat channels were detected automatically by MaxFilter, cross-checked manually and subsequently interpolated with the signal from neighbouring sensors. The data was further band-pass filtered between 0.2–40 Hz (zero-phase FIR filter), blink and heart artefacts were removed with the automated independent component analysis implemented in MNE Python. Components were rejected automatically based on high correlation value with the EOG and ECG electrodes. No baseline correction was performed. For the following mTRF (multivariate Temporal Response Function) analysis data was cut to include only speech segments (inter-stimulus intervals, breaks, and question responses were removed) and decimated to 200 Hz (no re-sampling was done to avoid anti-aliasing artefacts). Prior to joining the segments of the MEG data corresponding to speech segments, each segment was zero-padded along the time dimension to avoid discontinuous values and artificial jumps in signal amplitude. Only Gradiometer channels were used for the following analysis. Consistent with previous implementations of the mTRF analysis on electrophysiological data, we normalised (z-scored) data for each gradiometer channel thus converting the resulting mTRF weights into arbitrary units (Broderick et al., 2021).

2.3.5. mTRF analysis

mTRF analysis was conducted using the mTRF-Toolbox, an open-source MATLAB toolbox (Crosse et al., 2016). We fit forward encoding models using a multivariate feature set composed of the predictor vectors: Word Onsets, Speech Envelope and Semantic Surprisal. Word Onsets vectors consisted of spike functions with a single value (value 1) representing the onset of each word of each sentence in time. This vector was included to model out large evoked responses generated by word onsets and to minimise their impact on the estimates for the regressors of interest (Envelope and Semantic Surprisal). Speech envelopes were extracted from each spoken sentence using the same toolbox (*mTRFenvelope* function) which resampled the speech signal to match the sampling rate of the MEG data (200 Hz). The envelopes were then computed as the logarithmically scaled RMS intensity (square root of the average square of the nearest neighbours at every sample - Lalor and Foxe, 2010). Finally, the Semantic Surprisal vector was a scaled version of the Word onset vector, where vector height was equal to the negative log of the word probability given the preceding sentence context. These probabilities were derived using a custom-built neural network trained on a large set of online Spanish written corpora with the goal of predicting the next word in the sentence (see section below for details). As with MEG responses, vectors were zero-padded to ensure the same dimensionality as the MEG data.

All predictor vectors were simultaneously inputted to the mTRF analysis, which implements a lagged ridge regression describing a linear transformation of a given set of stimulus features $S(t,f)$ to the neural response $R(t,f)$ across a given set of lags (time shifts): $R(t,f) = mTRF* S(t,f)$. This analysis is run for each subject separately. Output of the mTRF is a set of weights for each stimulus feature f generated by the ridge regression where the ridge parameter was tuned across the folds of data to control for overfitting (Crosse et al., 2016). To estimate these weights, data was split into 10 folds of equal length (both MEG data and the stimulus vector). The model is trained on 9 of these folds and tested on the remaining fold - the resulting r value shows how well the model predicts the withheld data. Then the procedure is cross-validated across folds by leaving each fold out at a time. mTRF weights were generated for each

segment of the continuous data split into 10 folds and the resulting sets of weights were averaged. In this way we produced a unique set of mTRF weights across channels and lags for each subject.

To test if the resulting model could predict the neural responses at an above chance level and assess how accurately each feature of interest (Envelope and Semantic Surprisal) could capture the neural activity, for each subject we measured the model's ability to predict MEG responses above the null feature models (for a similar approach see Broderick et al., 2021). Two separate null models were derived for Envelope and Surprisal by randomly permuting the corresponding feature vector. During cross-validation testing the trained mTRF would attempt to predict neural responses from permuted features of interest while all other features were unchanged. Hence resulting neural signal predictions did not benefit from a given feature of interest. We repeated this step 10 times per feature vector and averaged the resulting null models' responses. We then subtracted the null model prediction accuracy (r value across sensors) from those of the true model. Surprisal and Envelope model accuracy correspond to the averages of these subtractions across the fronto-temporal sensors of each hemisphere. To test model accuracy at the group level we subjected these measures from all subjects to a one-sample t-test against zero. We found significant effects for both Envelope (LH $t = 22.4$, $p < 0.001$, RH $t = 21.1$, $p < 0.001$) and Surprisal (LH $t = 14.9$, $p < 0.001$, RH $t = 13.7$, $p < 0.001$) at the group level, showing that for the whole sample our full-feature model is better at predicting the neural signal than the null model.

2.3.6. Transformer neural network language model for Surprisal estimation

To estimate Semantic Surprisal (self-information) for every word in the sentence a self-attention-based autoregressive Transformer model (Vaswani et al., 2017) was used. That is, the language model considers previous words in a sentence to predict the next word. To optimise the model (lower memory consumption and increase training speed) two parameter-reduction techniques were applied: cross layer parameter sharing and factorization of parameters. The 15 M parameters model is based on 4 shared transformer layers with hidden layers of size 3072, embedding layer of size 768 (factorised with an intermediate layer of size 128) and 12 attention heads. The model was trained in one Nvidia Titan V GPU for one month, with a linear warm-up and scheduler and a learning rate peak set to 0.176e-3. We used mixed precision during training using a large batch size of 3.2 K sentences from the above corpora, excluding those sentences which were later used in the MEG experiment.

2.3.7. Grey matter analysis

A voxel-based morphometry (VBM) analysis was used to estimate grey matter volume in key areas of the reading network. For this analysis, individual T1-weighted images were preprocessed using the Statistical Parametric Mapping software (SPM12, Wellcome Department of Cognitive Neurology, London, UK) implemented in MATLAB (2014b, Mathworks, Inc, Natick, MA). First, the T1 images were reoriented and shifted to set the anterior commissure as the origin (0,0,0). Then, the images were segmented into grey matter, white matter, and cerebrospinal fluid (CSF) with the segmenting module implemented in SPM12. The volumes of the native segmentations were used to calculate the total intracranial volume (TIV) of each participant. Next, images were normalized into the Montreal Neurological Institute (MNI) space using the high-dimensional DARTEL normalization algorithm and further smoothed with an 8 mm full width half maximum (FWHM) Gaussian kernel. Finally, the processed images were used to estimate the volume features using the AAL atlas for brain parcellation. Specifically, we inverse normalized the AAL atlas to fit each participant's native space and selected a set of regions involved in the reading network (D'mello and Gabrieli, 2018), namely: inferior frontal gyrus (collapsing pars orbitalis, triangularis and opercularis), fusiform, superior temporal and temporo-parietal regions (collapsing angular gyrus and supramarginal gyrus). We considered regions bilaterally. Each volume of these regions of interest (ROI) was divided by the TIV to normalise the values.

2.4. Statistical analysis

The goal of the statistical analysis was to compare responses to auditory (Envelope) and contextual (Semantic Surprisal) features of the speech signal in the control and dyslexia groups and further to relate those measures to their reading abilities. The magnitude of the mTRF weights is a representation of the response to a given feature at a given channel and lag. To summarise responses to Envelope and Semantic Surprisal across channels and lags for each participant we (1) computed the global field power GFP (SD over channels) over fronto-temporal sensors and (2) computed the area under the curve for pre-defined lags of interest, referred to here as Energy. The mTRF model was fit with both Surprisal and Envelope within the same large window of lags (−100 to 700 ms) and output mTRF weights for Surprisal and Envelopes were further analysed in separate sets of lags. For Envelope response 0–200 ms lags were used since Envelope-specific mTRF responses were previously reported to peak within this lag range (Liberto et al. 2021). For Surprisal a large window of 300–700 ms lags was used, since responses to Surprisal using similar methods have been reported within this wider time range (Broderick et al., 2021).

First, to compare control and dyslexic groups directly, we used a two-sample matched t-test ($n = 14$ control and $n = 14$ dyslexic). To correlate mTRF response Energy with reading abilities we ran a multiple regression analysis on the whole larger sample ($n = 41$) with Surprisal and Envelope Energy as dependant variables, Word and Pseudoword reading scores as predictors and Age and non-verbal IQ (WAIS Matrix test) measure as nuisance regressors. We also entered a binary Group regressor (dyslexic coded as 1 and control as 0) and interaction terms for Group x Word and Group x Pseudoword reading into the model to test for differential relationship between mTRF responses and reading scores in dyslexia when compared with control group. We ran this regression in a hierarchical manner to determine whether Group and/or Group interactions significantly improved the variance explained by the model. We tested mTRF weights derived from the right and left hemisphere sensors separately. Identical regression analyses were also performed to correlate grey matter volume (GMV) with reading abilities, with GMV for each ROI (8 tested) as the dependant variable and Word, Pseudoword reading scores and mTRF Energy measures for Envelope and Surprisal as predictors, while Age and non-verbal IQ (WAIS Matrix test) measures were entered as nuisance regressors. Similarly, both a Group regressor and x Group interactions were added. We used false-discovery rate (FDR) correction to control for multiple comparisons across the 8 ROIs tested.

3. Results

3.1. mTRF results

We assessed the ability of our participants to track both auditory and semantic contextual features of the speech signal using multivariate Temporal Response Functions (mTRFs - see Methods). The input to this analysis is a multivariate vector, simultaneously including both the speech Envelope and Semantic Surprisal dimensions. The output of the mTRF analysis is a set of regression weights describing the relationship between each stimulus feature of interest and the MEG signal at each channel and time-lag. Larger weights represent a stronger contribution of a stimulus feature at a given time-lag to the MEG signal prediction. Fig. 1 shows both Envelope and Semantic Surprisal output weights (summarised over sensors as global field power GFP response) over time lags and corresponding topographies (averages over the time-lags). Dyslexic and Control matched groups are plotted separately for comparison. The peaks of the GFP show the temporal lags at which the MEG signal contributes most to the tracking of the corresponding function. Both Envelope and Semantic Surprisal were found to be significantly better predictors of the neural response within the fronto-temporal sensors compared to their respective null models (see Methods for details). Prediction ac-

curacy r for each feature was calculated as the difference between the prediction accuracy of the true and the null model (averaged within the fronto-temporal sensors of each hemisphere and across all subjects) and was as follows: Surprisal LH $r = 0.02$ and RH $r = 0.013$; Envelope LH $r = 0.033$ and RH $r = 0.035$ (see Supplementary Figure 1).

For Envelope tracking we observed an expected and previously reported (Di Liberto et al. 2021) response pattern - a peak at a 100 ms lag with largest responses in the bilateral temporal sensors for both groups. To test for an expected dyslexia-specific reduction of speech envelope tracking we conducted a one-sided t-test comparing GFP Energy - area under the curve within the 0–200 ms epoch - in the matched dyslexic/control sample. Consistent with previous literature reporting weaker cortical speech entrainment in dyslexia (Molinaro et al., 2016) we observed a reduced RH Envelope response in the Dyslexic group ($t = 1.83$, $p = 0.039$, one sided). No significant reduction of Envelope response was found in the LH ($t = 1.15$, $p = 0.129$).

For Surprisal tracking, the mTRF response pattern was also in line with previously reported results - peaks in responses within the 400 ms lag range in bilateral fronto-temporal sensors, corresponding to the well-documented N400 component sensitive to context effects (Broderick et al., 2018, 2021). Although responses in the dyslexic group were observably weaker than in the group of matched controls, these differences were not significant (t-test, GFP Energy - area under the curve - 300–700 ms epoch, LH $t = 1.57$, $p = 0.127$, RH $t = 0.89$, $p = 0.37$, two-sided, given no prior expectation about directionality of any effect).

Our main goal was to understand how tracking of perceptual and contextual information in speech is related to reading skills in both typical and dyslexic readers. Large variability in the dyslexic group´s reading skills (Dyslexic Word reading $M = 117.15$, SD = 24.3) showed that for some of this group, despite their overall worse phonological skills (Pseudoword reading $M = 66.64$, SD = 15.09), word reading skills approximated the lower end of the control group spectrum (Controls Word reading $M = 185.92$, SD = 37.8). Although, as expected, performance in reading words and pseudowords was related, the strength of this relation differed between groups: the correlation between Word and Pseudoword reading in the dyslexic sample was weaker ($r = 0.35$), compared to a stronger relationship ($r = 0.6$) in controls. We ran a hierarchical multiple regression analysis with GFP Energy (area under the curve) for Envelope and Surprisal as dependant variables and Word (WR), Pseudoword (PS) reading, non-verbal IQ (Matrix test) and Age as predictors in each hemisphere separately and selecting theoretically relevant fronto-temporal sensors (Klimovich-Gray et al., 2021). To understand how dyslexic readers differed from controls we asked if adding a group (control = 0, dyslexic = 1) regressor as well as interactions between group, word and pseudoword reading would significantly improve the model fit (F change).

This analysis revealed two main effects. First, stronger tracking of Envelope in the RH was marginally related to better pseudoword reading (PS $b = 0.3$, $p = 0.027$, model $F = 2.58$, $p = 0.054$) for all participants irrespective of Age ($b=−0.084$, $p = 0.75$) or non-verbal IQ ($b=−0.98$, $p = 0.186$). Adding a group regressor and related interactions did not significantly improve the model fit (F Change = 0.66, $p = 0.58$), suggesting this trend was similar for both controls and dyslexic groups. Second, stronger Surprisal response in the RH was related to better word reading skills in the dyslexic group (group by WR interaction $b = 0.52$, $p = 0.013$; F Change = 3.07, $p = 0.042$), although the relationship was non-significant for the whole sample (reduced model $F = 1.56$, $p = 0.2$). Fig. 2 (B) shows this clearly, with a positive correlation for dyslexia group but no apparent effect for controls. There is a similar but not significant trend of Surprisal in the LH. To contrast LH and RH effects directly, we compared regression beta values for the group by WR interaction in the two hemispheres using a bootstrap method to estimate 95% beta value confidence intervals (CIs). High overlap in the estimated beta CIs would mean that the strength of this effect is similar across both hemispheres. Given that we indeed found a high degree of overlap between the corresponding CIs, we conclude similar processes take place
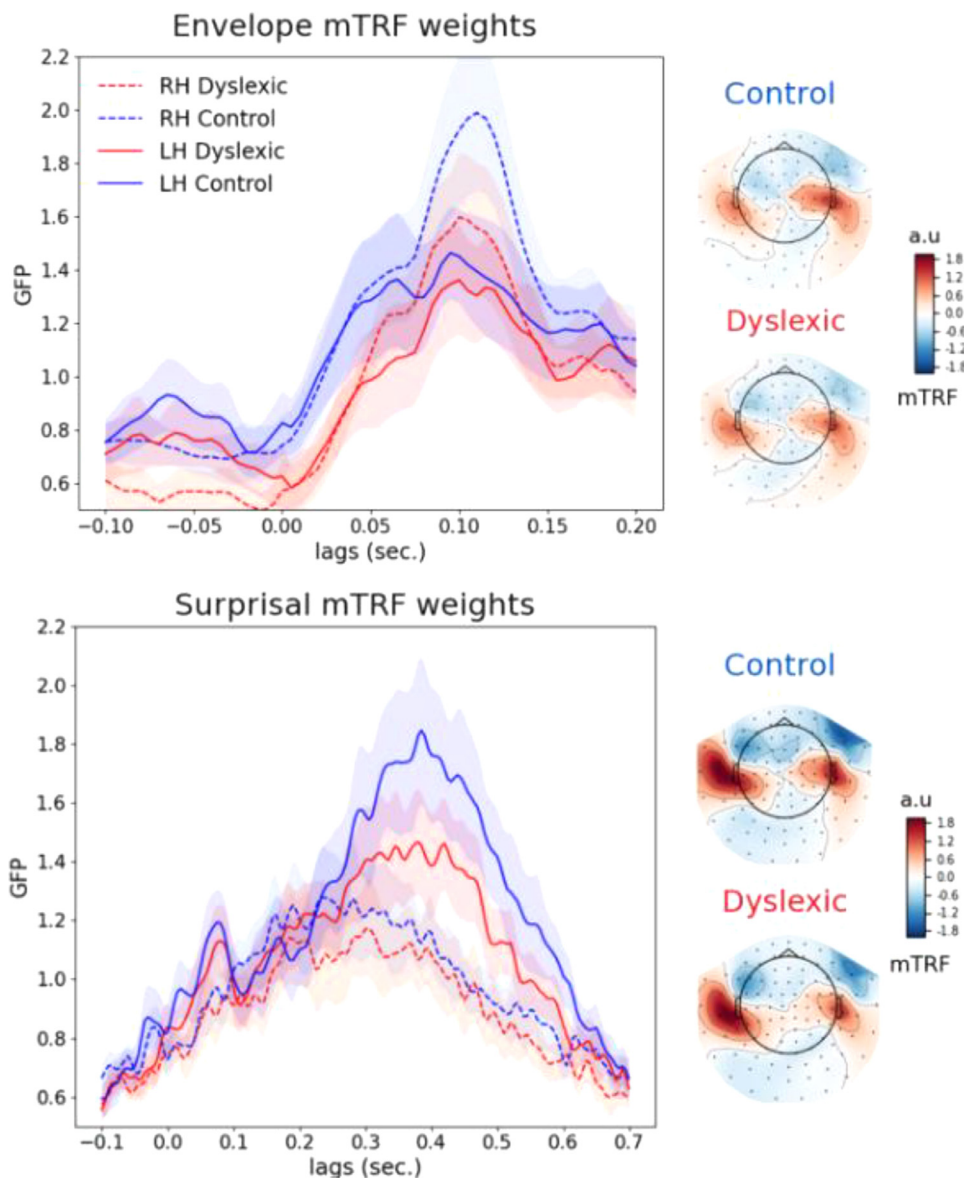
Fig. 1. mTRF weights showing the strength of the encoding of each feature within the neural signal. For Envelope (top) and Surprisal (bottom) the line plots (left) show the weights summarised using global field power (GFP, y axis) over the fronto-temporal sensors and plotted across the time lags of interest (x axis). LH and RH hemispheres are plotted separately with solid (LH) and dotted lines (RH). Matched controls are plotted in blue and dyslexic readers in red. Shaded areas show standard errors across subjects. The right panel shows corresponding sensor topographies which were averaged across all lags excluding the baseline. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in both hemispheres. Finally, Fig. 2 (C) shows in more detail differences in RH Surprisal tracking for worse and better dyslexic readers (group assigned based on the Word reading efficiency median split) - with better dyslexic readers showing a pattern of results comparable with the typical readers.

3.2. Grey matter volume analysis results

In the following analysis we asked if information tracking in speech (as measured by mTRF weights for Envelope and Surprisal) can be related to GMV in the areas of the reading circuit, as evidence of the link between speech processing and reading development. Given previous findings we expected that better reading (word and pseudoword) skills as well as better abilities in auditory and contextual information tracking in speech would be related to GMV increases in the bilateral frontal, temporal and parietal areas previously linked to reading proficiency. A preliminary analysis checking for potential differences in GMV across groups, showed that dyslexic and control readers had comparable GMV in each of the 8 ROIs (all p-values > 0.13; FDR corrected).

To address these questions, we ran multiple regressions with GMV as a dependant variable in 8 ROIs comprising the reading network

(D'mello and Gabrieli, 2018): bilateral superior temporal gyrus (STG), temporoparietal, fusiform and inferior frontal areas (see Fig. 3 for ROIs tested). As predictors we included the MEG mTRF variables that showed a significant relationship with reading skills: RH Envelope Energy and RH Surprisal Energy. To replicate previous work, we also included word and pseudoword reading scores and Age and non-verbal IQ as regressors (for full details of the GMV analysis see the Methods section). We also tested whether adding a group regressor and by group interactions (RH Envelope x Group and RH Surprisal x Group) would yield additional explanatory power to the model. We corrected for the number of ROIs tested using the FDR correction. This analysis was performed on a reduced subsample of controls (n = 20) for whom T1 MRI images were available. For the dyslexia group (n = 13) only one participant was removed due to a missing structural scan. We found a significant model fit in RH STG (model p = 0.002 unc.; p = 0.01, FDR corr.; R^2 = 0.42) and LH STG (model p = 0.006 unc.; p = 0.012, FDR corr.; R^2 = 0.37), RH IFG (model p = 0.004 unc.; p = 0.01, FDR corr.; R^2 = 0.38), LH IFG (model p = 0.004 unc.; p = 0.01, FDR corr.; R^2 = 0.39) and RH Fusiform gyrus (model p = 0.024 unc. p = 0.038, FDR corr.; R^2 = 0.28). Furthermore, we performed a "robustness check", and tested to what extent mTRF variables had an explanatory impact in the GMV mod-
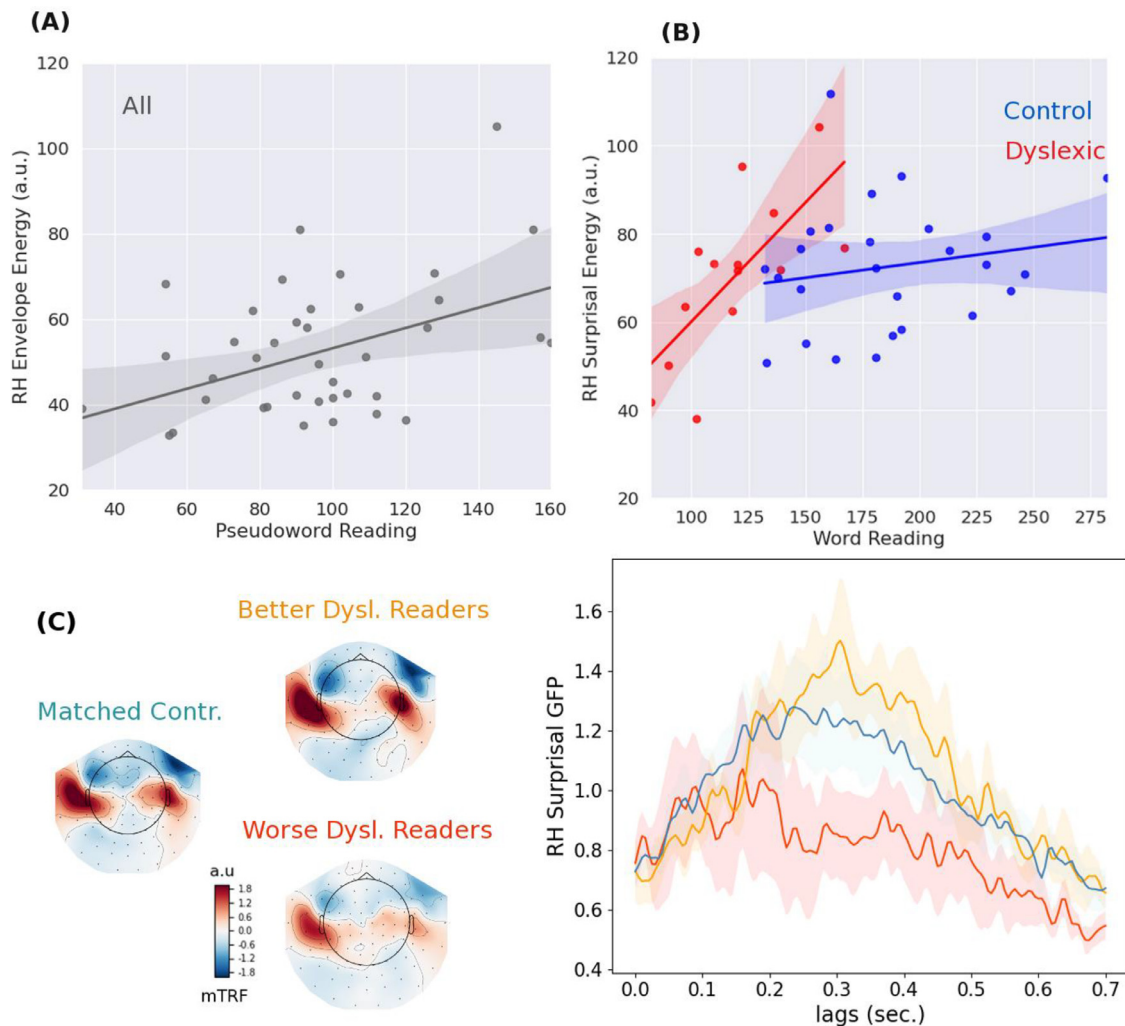
Fig. 2. Regression analysis. (A) A correlation plot with RH Envelope Energy (area under the GFP curve for the RH frontotemporal sensors) on the y axis and Pseudoword Reading scores on the x axis. A significant positive relationship between these two measures was present for the whole sample (dyslexic and controls, both plotted in grey). (B) A correlation plot with RH Surprisal Energy on the y axis and Word Reading scores on the x axis. A significant relationship between these two measures emerged only for dyslexic readers (in red) but not for controls (in blue). (C) mTRF weights for RH Surprisal plotted against time lags (right) and corresponding lag-averaged sensor topographies (left), plotted separately for matched controls in blue, worse dyslexic readers in red and better dyslexic readers in orange (better/worse median split based on word reading scores). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

els. In order to do so, we ran null models for each of the 8 selected ROIs leaving out the mTRF variables and just keeping word and pseudoword reading scores, Age, non-verbal IQ and group as regressors. None of the models reached significance (all p-values > 0.2; FDR corrected). This supports the idea that Envelope or Surprisal is related to GMV in those ROIs. Finally, adding group and by group interactions as regressors did not increase the explanatory power of any of the models.

RH Envelope Energy was related to GMV of the RH STG ($b = 0.64$, $t = 3.76$ $p <$ 0.001) and LH STG ($b = 0.4$, $t = 2.24$ $p = 0.03$) with better tracking being related to increased GMV. This was the same for both RH ($b = 0.58$, $t = 3.31$, $p = 0.003$) and LH ($b = 0.63$, $t = 3.62$, $p = 0.001$) IFG. While for the STG areas RH Envelope tracking was the only regressor that significantly contributed to the model fit, for both RH and LH IFG, word reading skills were also positively related to GMV ($b = 0.62$, $t = 2.62$, $p = 0.015$ for RH and $b = 0.64$, $t = 2.72$, $p = 0.012$ for LH). Despite the model showing a significant fit for RH Fusiform gyrus, no individual regressors were significant, so we cannot interpret which feature of neural or behavioural response was related to the GMV in this area.

In summary, GMV analysis revealed that better real-time tracking of auditory speech features (RH Envelope Energy) was related to increased GMV in areas related to auditory analysis (RH and LH STG) and those associated with higher-order language functions (RH and LH IFG). Moreover, in the bilateral IFG areas GMV was also positively correlated with Word reading skills.

4. Discussion

Early models proposed that dyslexic readers use context-based predictions to compensate for deficits in auditory and phonological processing. We used a novel combination of neuroimaging and machine learning techniques to gain insight into related neural mechanisms. We asked if the ability to track both auditory features of the speech envelope and contextual probabilities of words are related to reading skills in dyslexic and typical readers. We present three theoretically relevant findings. First, in line with previous proposals, we show that dyslexic readers are consistently worse at auditory envelope tracking and this further correlates with their phonological skills. Second, we show that the ability of participants with dyslexic symptoms to track contextual
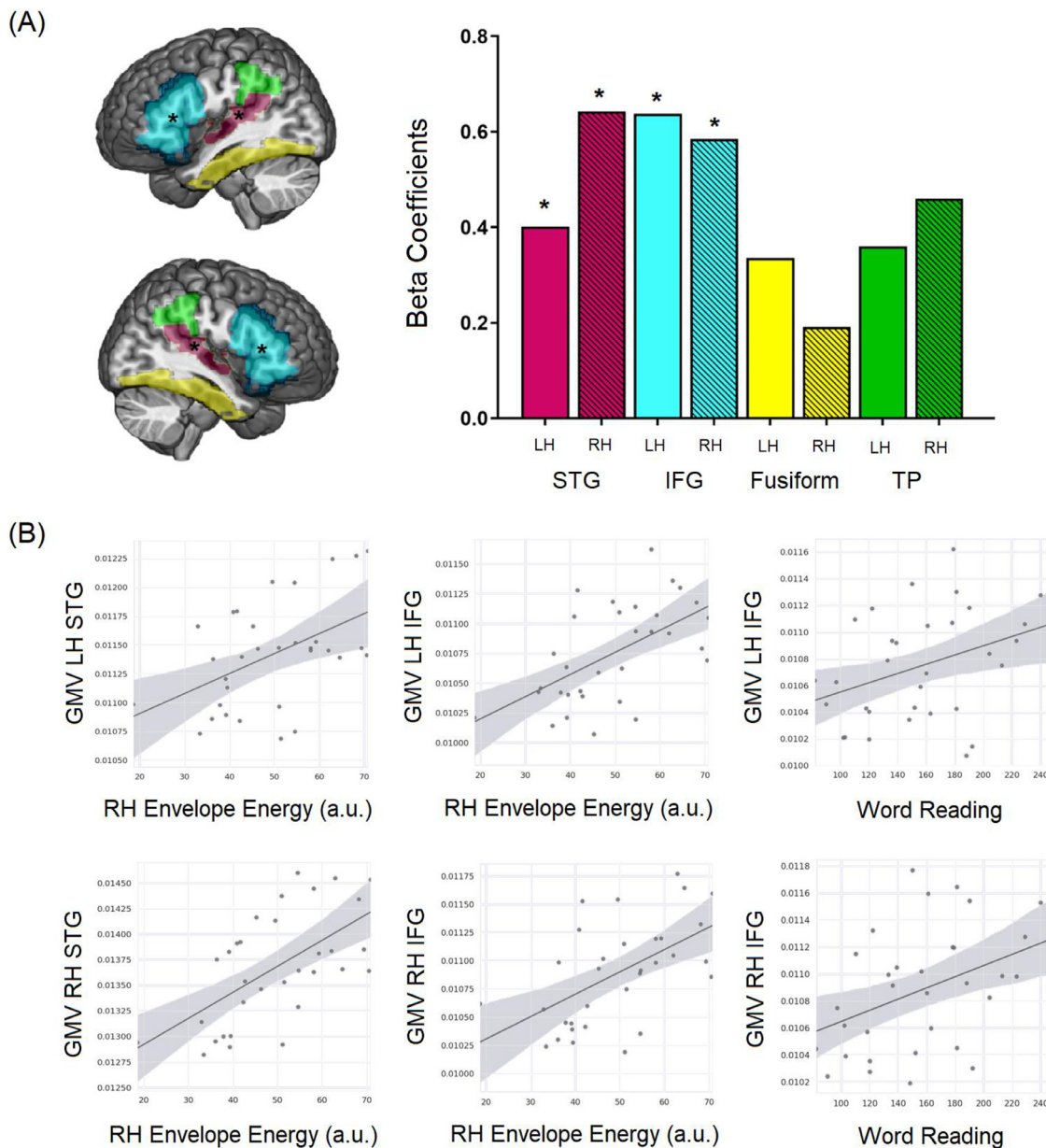
Fig. 3. (A) Relationship between online RH Envelope tracking captured with MEG (RH Envelope Energy derived from the mTRF analysis) and grey matter volume (GMV) in the ROIs linked to reading and speech processing. On the left are all eight bilateral ROIs that were tested, plotted in different colours on the template cortical surface. On the right is a bar chart where the height of each bar is a beta coefficient, showing how well RH Envelope tracking predicted GMV for a given ROI – significant effects are marked with a star *. RH ROIs are coded with a stripe pattern. (B) Scatterplots depicting significant associations between RH Envelope Energy and Word Reading with GMV in the STG and the IFG. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

probabilities in spoken language is linked to better word reading. Finally, we provide preliminary evidence that better Envelope tracking of speech is related to greater grey matter volume within the fronto-temporal areas (superior temporal and inferior frontal areas).

4.1. Envelope tracking linked to phonological skills

The speech envelope is dominated by lower frequencies (<10 Hz) that encode prosodic and syllabic speech features critical for detection of meaningful linguistic units (syllabic, word and phrase edges) and speech signal segmentation (Giraud and Poeppel, 2012; Lakatos, 2005; Schroeder and Lakatos, 2009). A growing literature has shown a replicable reduction of entrainment to the lower frequencies of the speech envelope in dyslexic readers (Lizarazu et al., 2015; Power et al., 2016;

Molinaro et al., 2016; but see Lizarazu et al., 2021 for related issues). This reduced synchrony with the speech envelope is thought to causally contribute to poor speech signal segmentation and phonological deficits and poor reading acquisition via the phonological route (Goswami, 2011).

Consistent with this, we show reduced RH speech envelope tracking in dyslexia and a positive relationship between Envelope tracking and phonological skills (measured by pseudoword reading) in the whole sample. This is consistent with recent work showing that both dyslexic adults and children have worse RH low-frequency (delta band <5 Hz) envelope (Molinaro et al., 2016) and phonological feature tracking (Di Liberto et al., 2018). These deficiencies could result from genetically related microstructural abnormalities of the auditory cortex found in dyslexic individuals (Giraud and Ramus, 2013), arguably contributing to

atypical hemispheric lateralisation in speech processing (Lehongre et al., 2011). Stronger RH Envelope tracking in controls is also consistent with the asymmetric sampling model (Abrams et al., 2008; Giraud and Poeppel, 2012; Poeppel, 2003), which suggests that due to cytoarchitectonic and microstructural morphology RH auditory areas are preferentially tuned to slower (delta-theta) aspects of the speech signal whereas homologous LH areas are tuned to faster (gamma) frequency ranges. Therefore, in line with previous work, we argue that RH Envelope tracking plays an important role in speech feature analysis by facilitating speech segmentation and its disruption in dyslexia has downstream effects for reading acquisition. Furthermore, here we show that Envelope tracking is related to phonological skills (pseudoword reading) across the whole sample, which implies it is also critical for typically developing readers.

Our exploratory GMV analysis is in line with this interpretation. If tracking of auditory envelopes in speech has a long-lasting impact on reading skills, we would expect a positive correlation between metrics of online performance and GMV within the ROIs that are part of the language and reading network. To our knowledge, this was the first attempt to relate neural measures of both Envelope and Semantic Surprisal to an individual's grey matter morphology. We observed that RH Envelope tracking was positively related with grey matter volume in bilateral STG and IFG, with greater volume of the IFG areas also being related to better word reading. The spatial distribution of these effects is in line with previous literature indicating bilateral STG and IFG as cortical sources of online speech Envelope tracking (vander Ghinst et al., 2016; Kubanek et al., 2013a). These studies also highlighted a functional differentiation between temporal and inferior frontal areas. STG Envelope tracking was not speech-specific and was thought to reflect tracking of signal acoustic properties (Kubanek et al., 2013b; Nourski et al., 2009). In contrast, speech-selective IFG Envelope tracking was shown to causally modulate oscillatory dynamics in the temporal areas (Di Liberto et al., 2018; Park et al., 2015) and was hypothesised to encode information about higher-order linguistic features (lexical/semantic/syntactic). In this context, the finding that bilateral IFG volume was predicted by both Envelope tracking and word reading skills suggests its involvement in both the auditory and linguistic feature analysis required for reading. Our combined GMV-mTRF analysis thus provides interesting preliminary findings that should be further validated and replicated.

The absence of significant effects in the STG GMV for word and pseudoword reading as well as the lack of group differences (controls vs dyslexia) has been discussed in a recent meta-analysis by Ramus et al. (2018), see also Torre and Eden (2019). A possible explanation is that RH Envelope tracking is a better predictor of the variance in the STG GMV compared to reading or grouping variables, hence when entered into the same model they explain no additional variance. If so, future studies of GMV and reading should consider lower-level auditory and perceptual processing measures alongside higher-order reading skills. While results in the dyslexia group should be interpreted with caution due to sample size (number of participants with dyslexia symptoms $n = 14$), previous studies have found effects of reduced speech Envelope tracking with comparable numbers (Molinaro et al., 2016; Power et al., 2016). Future work should be aimed at replicating these effects in a larger dyslexia sample and investigating similar effects in poor readers without dyslexia.

4.2. Semantic Surprisal tracking is related to better word reading in dyslexia

Dyslexic and typical readers did not differ in semantic Surprisal tracking. For both groups Surprisal response peaked around 400 ms post word onset, consistent with the N400 component latency linked to semantic contextual expectations (van Berkum et al., 2005; Kutas and Federmeier, 2011). EEG studies likewise have shown normal N400 responses in dyslexia for both spoken (Jednoróg et al., 2010; Rasamimanana et al., 2020) and written words in context (Rüsseler et al., 2007). In several studies N400 responses were delayed

or lasted longer in individuals with dyslexia (Jednoróg et al., 2010; Rüsseler et al., 2007) but here we find no such differences.

Critically, we show that Semantic Surprisal tracking was related to better word reading only in the dyslexic group (Fig. 2C). This is consistent with a seminal behavioural study by Nation and Snowling (1998) showing that dyslexic readers with good comprehension skills show stronger contextual facilitation in word reading. Although neither the previous nor current study designs allow us to make claims about the causal link between semantic context analysis and word reading in dyslexia, there are several plausible mechanisms. First, stronger semantic contextual analysis may coincide with better word reading for those with dyslexic symptoms who have milder reading deficits. However, were this the case, we would also expect those dyslexic readers to simultaneously show better phonological skills and for there to be a link between semantic context analysis and phonological processing skills. This was not the case in the current study - pseudoword reading did not predict the strength of cortical tracking of the semantic context. The other plausible mechanism is that deeper analysis of semantic relations facilitates lexical access when processing both spoken and written text without directly facilitating phonological skills but rather compensating for phonological deficits. In support of the latter, a recent study by Van der Kleij et al. (2019) showed that during picture-word priming, semantic but not phonological priming effects predicted word and pseudoword reading in older children with dyslexia (mean age 12). Interestingly, van Rijthoven et al. (2018) showed that the semantic abilities of younger dyslexic children (9-year-olds) predicted word naming and pseudoword decoding via rapid automatic naming and phonological awareness. Younger children, who are less experienced readers, rely more on the non-lexical reading route (Coltheart et al., 2001) and for them phonological and semantic skills may be more strongly coupled, compared with older children and the adults in our study, who rely primarily on the lexical route. Jointly, these studies reveal a complex picture of how the relationship between contextual semantic and phonological skills changes dynamically over time and highlight the need to understand these processes from a longitudinal perspective. Finally, an important next step for future research is to better understand the relationship between top-down semantic effects (as measured by Semantic Surprisal tracking) and envelope tracking in dyslexic readers who have overcome or partially overcome their initial reading impairments. A tentative prediction that can be made from the current set of results is that better readers with history of dyslexia will have a weaker relationship between envelope tracking and higher-order lexical and semantic processing skills since they have developed compensatory mechanisms allowing them to rely less on envelope entrainment.

Finding a positive relationship between fronto-temporal Surprisal tracking and word reading in dyslexia has further theoretically important implications for the neuro-cognitive study of dyslexia. Many previous studies reported hyperactivation of the RH fronto-temporal and temporo-parietal areas in dyslexic readers (Cutini et al., 2016; Hoeft et al., 2011a; Waldie et al., 2013). A possible explanation for this hyper-activation is increased cognitive effort due to over-reliance on rote memory learning (Rezaie et al., 2011). Here, for the first time, we show a positive dependency between language-relevant RH responses in participants with dyslexic symptoms (contextual semantic information tracking) and word reading. We therefore argue that RH activity in dyslexia is not simply a signature of cognitive effort but likely reflects compensatory mechanisms.

4.3. Conclusions and directions for future research

This study represents a first attempt to uncover the neural mechanisms underlying contextual facilitation of lexical access in dyslexic readers by applying multivariate decoding techniques to MEG data and relating these effects to behavioural and structural cortical measurements. We show an important dependency between speech envelope tracking, grey matter volume and pseudoword reading, supporting the

notion that dyslexic reading is related to inefficient speech signal sampling. We further uncover a link between semantic context tracking and better word reading skills in dyslexia, consistent with the compensation hypothesis.

Although this is a promising start, we cannot make any claims about the developmental aspects or causality of our effects. Since we did not have an additional reading-age matched control group we cannot determine whether the link between Surprisal tracking and word reading is specific to dyslexia, or a mechanism employed by various poor and learner readers. It can, however, be argued that for our adult sample the differential effect of reading experience should not be as pronounced as it would with young beginner readers. Hence our results are more likely to be specific to the dyslexic sample. Furthermore, it is not clear if the contextual compensation strategy is directly transferred from speech to reading or if both skills co-emerge due to other factors, such as intensive reading practice or explicit intervention. To answer this and related questions a follow-up longitudinal design with separate control groups matched for either chronological age or literacy level is necessary.

Declaration of Competing Interest

The authors declare no competing financial interests

Credit authorship contribution statement

Anastasia Klimovich-Gray: Conceptualization, Methodology, Software, Investigation, Formal analysis, Data curation, Writing – original draft, Visualization, Funding acquisition. **Giovanni Di Liberto:** Software, Validation, Writing – review & editing. **Lucia Amoruso:** Methodology, Formal analysis, Writing – review & editing. **Ander Barrena:** Methodology, Software, Writing – review & editing. **Eneko Agirre:** Methodology, Writing – review & editing. **Nicola Molinaro:** Conceptualization, Supervision, Funding acquisition.

Data availability

The primary data for this experiment is stored on secured servers at BCBL and can be made available via a request to the Authors upon reaching a prior formal data sharing agreement with BCBL.

The neuroimaging and behavioural data have been collected and is stored at BCBL (Basque centre on Cognition, Brain and Language) servers. This data can be shared upon reaching a formal data sharing agreement between BCBL and the host institution of the researcher who requests the data access.

Code availability - the following data packages were used in the analysis:

1. MNE Python platform and analysis pipeline (Gramfort et al., 2014) mne 1.0.3 available at https://github.com/mne-tools/mne-python/tree/maint/1.0.
2. MaxFilter version 2.2 by Elekta Neuromag, Elekta (2006). Max-Filter User's Guide.
3. mTRF Matlab toolbox (Crosse et al., 2016) available at https://github.com/mickcrosse/mTRF-Toolbox.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120072.

References

Abrams, D.A., Nicol, T., Zecker, S., Kraus, N., 2008. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. J. Neurosci. 28, 3958–3965.

Ahveninen, J., Hämäläinen, M., Jääskeläinen, I.P., Ahlfors, S.P., Huang, S., Lin, F.-.H., Raij, T., Sams, M., Vasios, C.E., Belliveau, J.W., 2011. Attention-driven auditory cortex short-term plasticity helps segregate relevant sounds from noise. Proc. Natl. Acad. Sci. U S A.

Bojar, O., Buck, C., Federmann, C., Haddow, B., Koehn, P., Leveling, J., Monz, C., Pecina, P., Post, M., Saint-Amand, H. et al. 2014. Findings of the 2014 workshop on statistical machine translation. Proceedings of the Ninth Workshop on Statistical Machine Translation.

Bonte, M.L., Blomert, L., 2004. Developmental dyslexia: ERP correlates of anomalous phonological processing during spoken word recognition. Cognit. Brain Res. 21, 360–376.

Broderick, M.P., Anderson, A.J., di Liberto, G.M., Crosse, M.J., Lalor, E.C., 2018. Electrophysiological correlates of semantic dissimilarity reflect the comprehension of natural, narrative speech. Curr. Biol. 28, 803–809 e3.

Broderick, M.P., di Liberto, G.M., Anderson, A.J., Rofes, A., Lalor, E.C., 2021. Dissociable electrophysiological measures of natural language processing reveal differences in speech comprehension strategy in healthy ageing. Sci. Rep. 11, 1.

Cardellino, C. 2016. Spanish Billion Words Corpus and Embeddings. https://crscardellino.github.io/SBWCE/ (Last accessed Feb 2019).

Chiappe, P., Chiappe, D.L., Gottardo, A., 2004. Vocabulary, context, and speech perception among good and poor readers. Educ. Psychol. (Lond.) 24, 825–843.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., Ziegler, J., 2001. DRC: a dual route cascaded model of visual word recognition and reading aloud. Psychol. Rev. 108, 204–256.

Corkett, J.K., Parrila, R., 2008. Use of context in the word recognition process by adults with a significant history of reading difficulties. Ann. Dyslexia. 58, 139–161.

Crawford, J.R., Howell, D.C., 1998. Comparing an individual's test score against norms derived from small samples. Clin. Neuropsychol. 12 (4), 482–486.

Crosse, M.J., di Liberto, G.M., Bednar, A., Lalor, E.C., 2016. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Front. Hum. Neurosci. 10.

Cuetos F., Rodríguez B., Ruano E., Arribas D. (2014) PROLEC-R-Batería de evaluación de los procesos lectores, Revisada. TEA Ediciones.

Cutini, S., Szűcs, D., Mead, N., Huss, M., Goswami, U., 2016. Atypical right hemisphere response to slow temporal modulations in children with developmental dyslexia. Neuroimage 143, 40–49.

Di Liberto, G.M., Lalor, E.C., Millman, R.E., 2018a. Causal cortical dynamics of a predictive enhancement of speech intelligibility. Neuroimage 166, 247–258.

Di Liberto, G.M., Peter, V., Kalashnikova, M., Goswami, U., Burnham, D., Lalor, E.C., 2018b. Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. Neuroimage 175, 70–79.

Di Liberto, G.M., Jingping, N., Yeaton, J., Khalighinejad, B., Shamma, S.A., Mesgarani, N., 2021. Neural Representation of Linguistic Feature Hierarchy Reflects Second-Language Proficiency. Neuroimage 227.

D'mello, A.M., Gabrieli, J.D.E., 2018. Cognitive neuroscience of dyslexia. Lang. Speech Hear Serv. Sch. 49, 798–809.

Dole, M., Meunier, F., Hoen, M., 2013. Gray and white matter distribution in dyslexia: a VBM study of superior temporal gyrus asymmetry. PLoS ONE 8.

Donhauser, P.W., Baillet, S., 2019. Two distinct neural timescales for predictive speech processing. Neuron 105, 1–9.

Frith, U., Snowling, M., 1983. Reading for meaning and reading for sound in autistic and dyslexic children. Br. J. Dev. Psychol. 1, 329–342.

vander Ghinst, M., Bourguignon, M., Beeck, M., Wens, V., Marty, B., Hassid, S., Choufani, G., Jousmäki, V., Hari, R., Bogaert, P., Goldman, S., Tiège, S., 2016. Left superior temporal gyrus is coupled to attended speech in a cocktail-party auditory scene. J. Neurosci. 36, 1596–1606.

Gillis, M., Vanthornhout, J., Simon, J.Z., Francart, T., Brodbeck, C., 2021. Neural markers of speech comprehension: measuring EEG tracking of linguistic speech representations, controlling the speech acoustics. J. Neurosci. 41, 10316–10329.

Giraud, A.-.L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. 15, 511–517.

Giraud, A.L., Ramus, F., 2013. Neurogenetics and auditory processing in developmental dyslexia. Curr. Opin. Neurobiol. 23, 37–42.

Goswami, U., 2011. A temporal sampling framework for developmental dyslexia. Trends Cogn. Sci. (Regul. Ed.) 15, 3–10.

Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Parkkonen, L., Hämäläinen, M.S., 2014. MNE software for processing MEG and EEG data. Neuroimage 86, 446–460.

Hakonen, M., May, P.J.C., Jääskeläinen, I.P., Jokinen, E., Sams, M., Tiitinen, H., 2017. Predictive processing increases intelligibility of acoustically distorted speech: behavioral and neural correlates. Brain Behav. 7.

Hale, J., 2016. Information-theoretical complexity metrics. Language Linguistics Compass 10 (9), 397–412.

Hämäläinen, J.A., Rupp, A., Soltész, F., Szücs, D., Goswami, U., 2012. Reduced phase locking to slow amplitude modulation in adults with dyslexia: an MEG study. Neuroimage 59, 2952–2961 2012.

Helenius, P., Salmelin, R., Service, E., Connolly, J.F., 1999. Semantic cortical activation in dyslexic readers. J. Cogn. Neurosci. 11, 535–550.

Hervais-Adelman, A.G., Carlyon, R.P., Johnsrude, I.S., Davis, M.H., 2012. Brain regions recruited for the effortful comprehension of noise-vocoded words. Lang. Cogn. Process. 27, 1145–1166.

Hoeft, F., McCandliss, B.D., Black, J.M., Gantman, A., Zakerani, N., Hulme, C., Lyytinen, H., Whitfield-Gabrieli, S., Glover, G.H., Reiss, A.L., Gabrieli, J.D.E., 2011a. Neural systems predicting long-term outcome in dyslexia. Proc. Natl. Acad. Sci. U. S. A. 108, 361–366.

Hoeft, F., Ueno, T., Reiss, A.L., Meyler, A., Whitfield-Gabrieli, S., Glover, G.H., Keller, T.A., Kobayashi, N., Mazaika, P., Jo, B., Just, M.A., Gabrieli, J.D.E., 2007. Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. Behav. Neurosci. 121, 602–613.

Jednoróg, K., Marchewka, A., Tacikowski, P., Grabowska, A., 2010. Implicit phonological and semantic processing in children with developmental dyslexia: evidence from event-related potentials. Neuropsychologia 48, 2447–2457.

Khalighinejad, B., Herrero, J.L., Mehta, A.D., Mesgarani, N., Zuckerman, M.B., 2017. Adaptation of the human auditory cortex to changing background noise. Nat. Commun. 10.

van der Kleij, Sanne W, Margriet, A.G., Segers, E., Verhoeven, L., 2019. Enhanced Semantic Involvement during Word Recognition in Children with Dyslexia. J. Exp. Child Psychol. 178, 15–29.

Klimovich-Gray, A., Barrena, A., Agirre, E., Molinaro, N., 2021. One way or another: cortical language areas flexibly adapt processing strategies to perceptual and contextual properties of speech. Cerebral Cortex 31, 4092–4103.

Krafnick, A.J., Lynn Flowers, D., Luetje, M.M., Napoliello, E.M., Eden, G.F., 2014. An investigation into the origin of anatomical differences in dyslexia. J. Neurosci. 34, 901–908.

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., Schalk, G., 2013a. The tracking of speech envelope in the human cortex. PLoS ONE.

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., Schalk, G., 2013b. The tracking of speech envelope in the human cortex. PLoS ONE.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.

Lakatos, 2005. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J. Neurophys. 94, 1904–1911.

Lalor, E.C., Foxe, J.J., 2010. Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. Eur. J. Neurosci. 31, 189–193.

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., Giraud, A.L., 2011. Altered Low-Gamma Sampling in Auditory Cortex Accounts for the Three Main Facets of Dyslexia. Neuron 72, 1080–1090.

Lizarazu, M., Lallier, M., Molinaro, N., 2018. Desynchronization between speech rhythms and neural oscillations: a possible cause of phonological problems in dyslexia. Anales RANM 135, 47–51.

Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P.M., Lerma-Usabiaga, G., Carreiras, M., 2015. Developmental evaluation of atypical auditory sampling in dyslexia: functional and structural evidence. Hum. Brain Mapp. 36, 4986–5002.

Lizarazu, M., di Covella, L.S., van Wassenhove, V., Rivière, D., Mizzi, R., Lehongre, K., Hertz-Pannier, L., Ramus, F., 2021. Neural entrainment to speech and nonspeech in dyslexia: conceptual replication and extension of previous investigations. Cortex 137, 160–178.

Mody, M., Wehner, D.T., Ahlfors, S.P., 2008. Auditory word perception in sentence context in reading-disabled children. Neuroreport 19, 1567–1571.

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., Carreiras, M., 2016. Out-of-synchrony speech entrainment in developmental dyslexia. Hum. Brain Mapp. 37.

Nation, K., Snowling, M.J., 1998. Individual differences in contextual facilitation: evidence from dyslexia and poor reading comprehension. Child Dev. 69, 996–1011.

Nourski, K.V., Reale, R.A., Oya, H., Kawasaki, H., Kovach, C.K., Chen, H., Howard, M.A., Brugge, J.F., 2009. Temporal envelope of time-compressed speech represented in the human auditory cortex. J. Neurosci. 29, 15564–15574.

Padró L., Rigau G., Reese S., Boleda G., Cuadros M. (2010) Wikicorpus: a Word-Sense Disambiguated Multilingual Wikipedia Corpus. Wikicorpus: a Word-Sense Disambiguated Multilingual Wikipedia Corpus. Available at: http://www.fsf.org/licensing/licenses/fdl.html.

Park, H., Ince, R.A.A., Schyns, P.G., Thut, G., Gross, J., 2015. Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. Curr. Biol. 25, 1649–1653.

Peyre, H., Mohanpuria, N., Jednoróg, K., Heim, S., Grande, M., van Ermingen-Marbach, M., Altarelli, I., Monzalvo, K., Williams, C.M., Germanaud, D., Toro, R., 2020. Neuroanatomy of dyslexia: An allometric approach. Eur. J. Neurosci. 52, 3595–3609.

Poeppel D. (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time.' Speech Commun. 41:245–255.

Power, A.J., Colling, L.J., Mead, N., Barnes, L., Goswami, U., 2016. Neural encoding of the speech envelope by children with developmental dyslexia. Brain Lang. 160, 1–10.

Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., Di Covella, L.S., 2018. Neuroanatomy of developmental dyslexia: pitfalls and promise. Neurosci. Biobehav. Rev. 84, 434–452.

Rasamimanana, M., Barbaroux, M., Colé, P., Besson, M., 2020. Semantic compensation and novel word learning in university students with dyslexia. Neuropsychologia 139.

Rezaie, R., Simos, P.G., Fletcher, J.M., Cirino, P.T., Vaughn, S., Papanicolaou, A.C., 2011. Temporo-parietal brain activity as a longitudinal predictor of response to educational interventions among middle school struggling readers. J. Int. Neuropsychol. Soc. 17, 875–885.

van Rijthoven, R., Kleemans, T., Segers, S., Verhoeven, L., 2018. Beyond the phonological deficit: semantics contributes indirectly to decoding efficiency in children with dyslexia. Dyslexia 24, 309–321.

Rüsseler, J., Becker, P., Johannes, S., Münte, T.F., 2007a. Semantic, syntactic, and phonological processing of written words in adult developmental dyslexic readers: an event-related brain potential study. BMC Neurosci. 8.

Schäfer, R., Bildhauer, F., 2012. Building large corpora from the web using a new efficient tool chain. In: Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12).

Schroeder, C.E., Lakatos, P., 2009. Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. 32, 9–18.

Schulz, E., Maurer, U., van der Mark, S., Bucher, K., Brem, S., Martin, E., Brandeis, D., 2008. Impaired semantic processing during sentence reading in children with dyslexia: combined fMRI and ERP evidence. Neuroimage 41, 153–168.

Shaywitz, S.E., Fletcher, J.M., Holahan, J.M., Shneider, A.E., Marchione, K.E., Stuebing, K.K., Francis, D.J., Pugh, K.R., Shaywitz, B.A., 1999. Persistence of dyslexia: the connecticut longitudinal study at adolescence. Paediatrics 104, 1351–1359.

Silva-Pereyra, J., Rivera-Gaxiola, M., Fernández, T., Díaz-Comas, L., Harmony, T., Fernández-Bouzas, A., Rodríguez, M., Bernal, J., Marosi, E., 2003. Are poor readers semantically challenged? An event-related brain potential assessment. Int. J. Psychophysiol. 49, 187–199.

Sjerps, M.J., Fox, N.P., Johnson, K., Chang, E.F., 2019. Speaker-normalized sound representations in the human auditory cortex. Nat. Commun. 10, 1–9.

Stanovich, K.E., 1980. Toward an interactive-compensatory model of individual differences in the development of reading fluency. Read. Res. Q. 32–71.

Taulu, S., Simola, J., Kajola, M., 2005. Applications of the signal space separation method. Signal Processing 53 (9), 3359–3372 IEEE.

Tervaniemi, M., Putkinen, V., Nie, P., Wang, C., Du, B., Lu, J., Li, S., Cowley, B.U., Tammi, T., Tao, S., 2022. Improved auditory function caused by music versus foreign language training at school age: is there a difference? Cerebral Cortex 32, 63–75.

Torre, G.A.A., Eden, G.F., 2019. Relationships between gray matter volume and reading ability in typically developing children, adolescents, and young adults. Dev. Cogn. Neurosci. 36.

van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V., Hagoort, P., 2005. Anticipating Upcoming Words in Discourse: evidence From ERPs and Reading Times. J. Exp. Psychol. 31, 443–467.

Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser L., Polosukhin I. (2017) Attention Is All You Need. 31st Conference on Neural Information Processing Systems (NIPS 2017).

Waldie, K.E., Haigh, C.E., Badzakova-Trajkov, G., Buckley, J., Kirk, I.J., 2013. Reading the wrong way with the right hemisphere. Brain Sci. 3, 1060–1075.