Katarzyna Wasilewska

Functional Magnetic Resonance Spectroscopy (fMRS) Study on Reading-Related Metabolite Changes

PhD thesis

completed at the Laboratory of Language Neurobiology

of the Nencki Institute of Experimental Biology

Polish Academy of Sciences

Supervisor:

Prof. Katarzyna Jednoróg, Ph.D., D.Sc.

Assistant Supervisor:

Bartosz Kossowski, Ph.D.

Warsaw, 2025

This research was funded by the National Science Center OPUS grant (2019/35/B/HS6/01763) awarded to Prof. Katarzyna Jednoróg. The study was approved by the institutional review board at the Faculty of Psychology, University of Warsaw, Poland (reference number 2N/02/2021).

TABLE OF CONTENTS

Abstract

Functional Magnetic Resonance Spectroscopy (fMRS) is a non-invasive technique used to measure dynamic changes in metabolite concentrations in response to stimuli. Despite its potential for advancing our understanding of brain activation mechanisms, fMRS remains relatively novel and the temporal dynamics of glutamate (Glu), the main excitatory neurotransmitter, following stimulation have not yet been fully explored. To date, no studies have applied fMRS to the reading process, despite the potential of this approach to reveal dynamic glutamate responses that may underlie both typical reading and its impairments in dyslexia. One of the newest mechanistic account of dyslexia, the neural noise hypothesis, suggests that it could be caused by an imbalance between glutamate and gamma-aminobutyric acid (GABA), the main inhibitory neurotransmitter. In particular, an elevated concentration of glutamate in the left superior temporal sulcus (STS) was proposed to disrupt signal processing and impair reading acquisition.

The aim of this thesis was to investigate glutamate concentration changes during reading-related tasks, in brain regions involved in reading: the superior temporal sulcus and the visual word form area (VWFA), as well as in one control region, the medial prefrontal cortex (mPFC). To characterize the temporal dynamics of glutamate, fMRS signals were acquired at four different delays between stimulus onset and signal acquisition. Participants with varying reading abilities, including individuals diagnosed with dyslexia and typical readers, were scanned at both 7T and 3T MR scanners. In total, 59 participants (29 with dyslexia, 13 females; 30 typical readers, 14 females) were scanned at 7T, and 40 participants (21 with dyslexia, 9 females; 19 typical readers, 11 females) at 3T. Glutamate levels were compared between groups to determine whether participants diagnosed with dyslexia exhibit higher glutamate concentrations in reading-related brain regions. While 7T scanners theoretically provide higher spectra resolution and improved metabolite separation, they also introduce technical challenges.

For the VWFA, reliable analysis was not feasible due to insufficient spectral quality, highlighting the methodological difficulty of collecting data from regions susceptible to magnetic field inhomogeneities. In the STS, glutamate responses to reading-related stimulation were heterogeneous. Effects were more apparent in females, yet they were sensitive to blood oxygenated level depended (BOLD) correction and varied between 7T

and 3T. No evidence of elevated glutamate in dyslexic participants within the left STS was observed, which does not support the neural noise hypothesis. Glutamate concentration changes were not limited to reading-sensitive regions, and some responses were also observed in the mPFC. A consistent glutamate response function could not be established, as glutamate changes varied across sex, group, brain region, stimulation type, and scanner. This inconsistency may reflect limited spectral quality due to a small number of averaged signals and the impact of BOLD contamination. Additionally, glutamate levels were significantly influenced by sex, age, and voxel tissue composition. While 7T improved some quality parameters, overall gains over 3T were inconsistent and region-dependent. These findings suggest that the practical advantages of ultra-high-field scanners in fMRS depend on region and are constrained by technical challenges.

Streszczenie

Funkcjonalna spektroskopia rezonansu magnetycznego (fMRS) to nieinwazyjna metoda pozwalająca mierzyć dynamiczne zmiany stężeń metabolitów w mózgu w odpowiedzi na bodźce. Mimo dużego potencjału w badaniu mechanizmów aktywacji mózgu, fMRS pozostaje techniką relatywnie nową, a dynamika czasowa odpowiedzi glutaminianu (Glu), głównego neuroprzekaźnika pobudzającego, po stymulacji nie została jeszcze w pełni zbadana. Do tej pory żadne badania nie zastosowały fMRS w procesie czytania, mimo że podejście to mogłoby ujawnić dynamiczną odpowiedź glutaminianu leżącą u podstaw zarówno typowego czytania, jak i jego zaburzeń w dysleksji. Jedna z najnowszych hipotez dotyczących mózgowych mechanizmów dysleksji – hipoteza szumu neuronalnego – sugeruje, że może być ona spowodowana zaburzeniem równowagi między glutaminianem a kwasem gamma-aminomasłowym (GABA), głównym neuroprzekaźnikiem hamującym. W szczególności zaproponowano, że podwyższone stężenie glutaminianu w lewym górnym zakręcie skroniowym (STS) zakłóca przetwarzanie sygnałów i utrudnia nabywanie umiejętności czytania.

Celem rozprawy było zbadanie zmian stężenia glutaminianu podczas zadań związanych z czytaniem w obszarach mózgu zaangażowanych w czytanie: STS oraz w obszarze wzrokowej formy słów (VWFA) w lewej korze skroniowo-potylicznej, a także w rejonie kontrolnym, przyśrodkowej korze przedczołowej (mPFC). Aby uchwycić dynamikę zmian glutaminianu w czasie, sygnały fMRS pozyskiwano przy czterech różnych opóźnieniach po rozpoczęciu bodźca. Uczestnicy ze zróżnicowanym poziomem umiejętności czytania, w tym osoby z dysleksją i typowo czytające, byli badani na dwóch skanerach 7T i 3T. W sumie 59 osób (29 z dysleksją, 13 kobiet; 30 typowych czytelników, 14 kobiet) zbadano na skanerze 7T, a 40 osób (21 z dysleksją, 9 kobiet; 19 typowych czytelników, 11 kobiet) na skanerze 3T. Stężenia glutaminianu porównywano między grupami, aby sprawdzić, czy uczestnicy z dysleksją wykazują wyższe wartości w obszarach mózgu związanych z czytaniem. Choć skanery 7T teoretycznie zapewniają wyższą rozdzielczość częstotliwościową, co pozwala na dokładniejsze rozróżnienie metabolitów, wiążą się z dodatkowymi trudnościami technicznymi.

Ze względu na trudną lokalizację VWFA w obszarze narażonym na niejednorodności pola magnetycznego, nie udało się uzyskać wystarczającej liczby widm dobrej jakości,

by przeprowadzić analizy tego regionu. W STS odpowiedzi glutaminianu na bodźce związane z czytaniem były zróżnicowane. Efekty były bardziej widoczne u kobiet, ale w dużym stopniu zależały od zastosowania korekcji BOLD i różniły się między skanerami 7T i 3T. Wyniki nie wykazały wyższego poziomu glutaminianu w lewym STS u osób z dysleksją jak przewidziano w hipotezie szumu neuronalnego. Zmiany stężenia glutaminianu obserwowano nie tylko w obszarze związanym z czytaniem, lecz także w mPFC. Ze względu na zmienność wyników w zależności od płci, grupy, obszaru mózgu, rodzaju bodźca i skanera nie udało się ustalić funkcji zmian glutaminianu w czasie. Może to być spowodowane niewystarczającą jakością analizowanych widm powstałych z uśrednienia stosunkowo niewielkiej liczby sygnałów dla każdego z punktów czasowych oraz wpływem efektu BOLD. Zaobserwowano również, że płeć, wiek i skład tkankowy analizowanego woksela znacząco wpływały na poziomy glutaminianu. Niektóre parametry jakości widma były lepsze przy użyciu skanera 7T, jednak ogólne korzyści w porównaniu do 3T były niejednoznaczne i zależały od analizowanego obszaru mózgu.

Abbreviations

ACC	anterior cingulate cortex
ARHQ	Adult Reading History Questionnaire
BOLD	blood oxygenation level–dependent
Cho	choline
CRLB	Cramér-Rao lower bound
CSDE	chemical shift displacement error
CSF	cerebrospinal fluid
EEG	electroencephalography
FID	free induction decay
fMRI	functional magnetic resonance imaging
fMRS	functional magnetic resonance spectroscopy
FWHM	full width at half maximum
GABA	gamma-aminobutyric acid
GLM	general linear model
Gln	glutamine
Glu	glutamate
Glx	glutamate + glutamine
GM	gray matter
GRF	glutamate response function
GSH	glutathione
Hz	hertz
IFG	inferior frontal gyrus
Lac	lactate
LSS	letter–speech sound
mI	myo-Inositol
MM	macromolecules
mM	millimolar
mPFC	medial prefrontal cortex
MRI	magnetic resonance imaging
MRS	magnetic resonance spectroscopy
MRSI	magnetic resonance spectroscopic imaging
NAA	N-acetylaspartate

ppm	parts per million
PRESS	Point RESolved Spectroscopy
RAN	rapid automatized naming
RF	radiofrequency
SAR	specific absorption rate
SNR	signal-to-noise ratio
STEAM	STimulated Echo Acquisition Mode
STG	superior temporal gyrus
STS	superior temporal sulcus
tCr	total creatine
TE	echo time
TR	repetition time
VWFA	visual word form area
WM	white matter

1. Introduction

1.1. Magnetic resonance spectroscopy (MRS)

Magnetic resonance spectroscopy (MRS) is a method used to analyze the chemical composition of selected tissues or materials. This non-invasive technique utilizes clinical magnetic resonance scanners, which are widely available in hospitals and research facilities, to identify and quantify metabolites (molecules involved in the body's metabolism).

For atomic nuclei to be visible in MRS, they must have magnetic moment (i.e., a spin different from zero), allowing them to interact with the magnetic field. Additionally, a high gyromagnetic ratio (unique to each nucleus) allows for high sensitivity in measurements. Furthermore, the isotope should have a high natural abundance in the body. Three isotopes are considered the most suitable for MRS experiments: phosphorus (^{31}P: 15 protons, 16 neutrons; $spin = {}^{1}/_{2}$, gyromagnetic ratio = 17.24 MHz/T), carbon (^{13}C: 6 protons, 7 neutrons; $spin = {}^{1}/_{2}$, gyromagnetic ratio = 10.71 MHz/T), and hydrogen (^{1}H: 1 proton, 0 neutrons; $spin = {}^{1}/_{2}$, gyromagnetic ratio = 42.58 MHz/T).

Phosphorus magnetic resonance spectroscopy (^{31}P-MRS) offers valuable insights into energy metabolism by detecting high-energy metabolites and membrane phospholipids. However, its clinical application is limited due to its relatively low sensitivity and the low concentrations of these metabolites, which result in a low signal-to-noise ratio (SNR), coarse spatial resolution, and prolonged acquisition times (Santos-Díaz & Noseworthy, 2020; Andrade et al., 2014).

Carbon MRS (^{13}C-MRS) allows for monitoring metabolic flux and neurochemical dynamics, offering valuable insights into energy production and neurotransmitter activity. However, its clinical use is limited by low sensitivity, partly due to the natural abundance of ^{13}C is only about 1.1% of all carbon isotopes, as well as poor spatial resolution and long acquisition times. Despite these challenges, technological advancements are improving signal quality and acquisition speed (Gruetter et al., 2003; Ross et al., 2003).

Hydrogen (^{1}H-MRS, proton) is most commonly used in MRS due to its strong magnetic moment, high sensitivity, and prevalence in the human body. The first reports describing proton magnetic resonance spectroscopy of the brain, including the measurement of N-acetylaspartate (NAA) using a 1.5 Tesla scanner and the STEAM sequence, were published in 1989 (Narayana et al., 1989). This marked the beginning of the rapid development of the MRS technique. Single-voxel spectroscopy, which measures chemical compounds in a selected brain region (a voxel), has since become widely used, not only by researchers but also in clinical settings to assist with diagnoses. Nevertheless, its clinical application remains limited.

The MRS method could potentially be used in the future to non-invasively grade brain tumors instead of histopathological diagnosis, which requires a biopsy. There are studies suggesting that the Cho/Cr and Cho/NAA ratios may serve as useful markers in distinguishing between low- and high-grade gliomas (Shokry, 2012). Despite its valuable role in diagnosis and treatment, its diagnostic accuracy can be limited because different pathologies may share overlapping metabolic profiles, which complicates differential diagnosis (Weinberg et al., 2021). Nevertheless, a recent study showed that using full metabolite profiles together with machine learning enables 93% diagnostic accuracy in differentiating medulloblastoma, pilocytic astrocytoma, and ependymoma (Gill et al., 2024). Magnetic resonance spectroscopy is also widely used to investigate the neurobiological basis of psychiatric disorders. For instance, Duarte and Xin (2019) reported an imbalance in the glutamine-to-glutamate (Gln/Glu) ratio in patients with schizophrenia, which may indicate disturbances in the glutamatergic cycle, along with reduced levels of N-acetylaspartate (NAA). In addition to brain research, MRS is also used to analyze the chemical composition of the breast (Begley et al., 2012; Baltzer & Dietzel, 2013), prostate (Stamatelatou et al., 2022; Bellomo et al., 2016), liver (Pasanta et al., 2021; Hamilton et al., 2009), muscle (Xu et al., 2012; Deshmukh et al., 2014), and spinal cord (Wyss et al., 2017).

Another important technique based on the phenomenon of magnetic resonance spectroscopy is magnetic resonance spectroscopy imaging (MRSI). This multi-voxel technique utilizes phase-encoding to acquire MRS data from many voxels, enabling the measurement of the spatial distribution of metabolite concentrations (Skoch et al., 2008; Posse et al., 2013; Maudsley et al., 2021).

Moreover, the new technique of functional magnetic resonance spectroscopy (fMRS) has recently become more attractive. fMRS enables the study of neurotransmitter concentration changes in the brain in response to stimuli. This novel method offers potential for improving our understanding of how the brain responds to various types of stimulation (Stanley & Raz, 2018).

1.1.1. Physical principle of ^{1}H-MRS

Proton magnetic resonance spectroscopy (^{1}H-MRS) is based on the same physical principles as magnetic resonance imaging (MRI). In both techniques, signals from hydrogen protons are used either to generate anatomical images (MRI) or to obtain a frequency spectrum of biochemical compounds (MRS). Since protons are most abundant in water, which is present at concentrations approximately 10^4 times higher than other metabolites, water suppression is necessary to detect the much weaker metabolite signals (Tkáč et al., 2021). Excited hydrogen nuclei generate a signal in the time domain, which is then converted into the frequency domain using the Fourier Transform.

Protons in different chemical environments experience slightly different local magnetic fields. This variation arises from the motion of electrons surrounding the nucleus, which generate small circulating currents. These currents produce secondary magnetic fields that oppose the main magnetic field (B_0), a phenomenon known as electron shielding. As a result, protons resonate at slightly different frequencies depending on their chemical environment.

The resonance frequency of a nucleus, known as the Larmor frequency, depends on the gyromagnetic ratio (γ) and the magnetic field induction (B_0):

$$f_{Larmor} = \gamma \cdot B_0$$

Due to shielding, the effective resonance frequency is slightly shifted and described by:

$$f_{resonance} = \gamma \cdot B_0 \cdot (1 - \sigma)$$

where:

γ – gyromagnetic ratio (for protons: 42.58 MHz/T)

B_0 – magnetic field induction

σ – shielding constant (depends on chemical environment)

This shift allows us to distinguish signals from protons in different molecules, as they appear at distinct positions along the frequency axis of the spectrum (Tognarelli et al., 2015). Frequency shifts, known as chemical shifts *(*δ), are expressed in Hertz (Hz) and are directly proportional to the magnetic field induction (B_0). To standardize spectra acquired on scanners with different field strengths, chemical shifts are typically reported in parts per million (ppm):

$$\delta\ [\text{Hz}] = \nu_{sample} - \nu_{ref}$$

$$\delta\ [\text{ppm}] = \frac{\nu_{sample} - \nu_{ref}}{\nu_{ref}} \cdot 10^6$$

where:

δ – chemical shift

ν_{sample} – resonance frequency of the measured proton [Hz]

ν_{ref} – reference frequency (typically tetramethylsilane, TMS)

Historically, tetramethylsilane ($Si(CH_3)_4$) has been used as a reference standard for proton MRS (Harris et al., 2002). Because of this normalization, peaks of specific metabolites consistently appear at the same ppm values regardless of the magnetic field strength e.g., NAA at ~2.02 ppm and water at ~4.68 ppm.

Higher magnetic field strengths improve spectral resolution, making it easier to resolve overlapping peaks. Additionally, the resonance of one proton can be affected by neighboring nuclei. When nearby hydrogen nuclei interact, the resulting spin-spin coupling leads to splitting of the spectral lines into multiplets, a phenomenon known as J-coupling (Faghihi et al., 2017). Each metabolite has a characteristic spectral "fingerprint," with signals at specific frequencies and multiplet patterns. The MRS spectrum represents the sum of these individual spectral signatures, weighted by the concentration of each metabolite in the tissue.

We are able to observe a range of brain metabolites using ^{1}H-MRS. The most prominent ones are shown in Figure 1, which presents a representative spectrum with annotated peak positions.

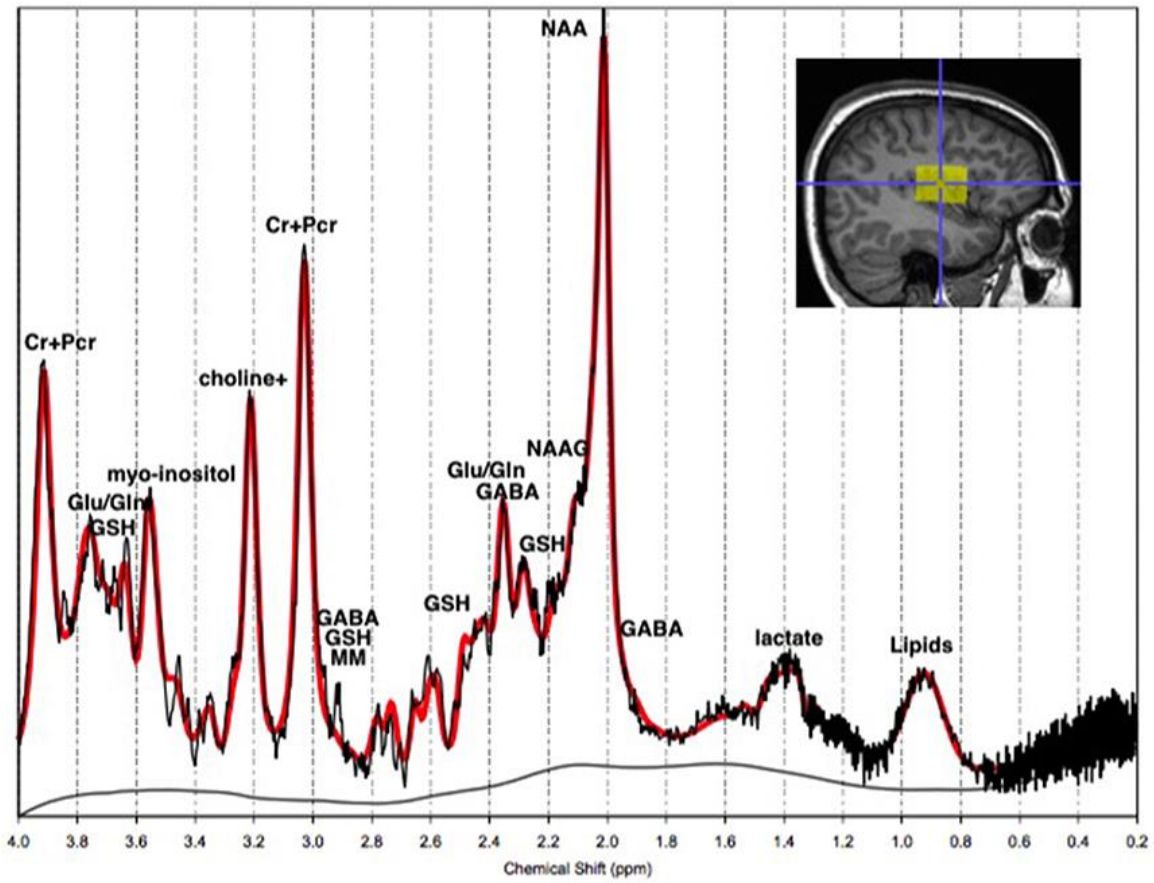

Figure 1. Spectrum reprinted from A Comprehensive Review of the ^{1}H-MRS Metabolite Spectrum in Autism Spectrum Disorder (Ford & Crewther, 2016). Licensed under CC BY.

The metabolite concentrations listed in Table 1 represent approximate physiological ranges observed in the adult human brain. These values have been rounded for clarity and are based on Govindaraju et al. (2000) and the ISMRS lecture by Ulrike Dydak (2023).

Table 1. Representative brain metabolites detectable with ^{1}H-MRS **(**The concentrations of metabolites are calculated as area under the peak corresponding to this metabolite).

Metabolite	**Function**	**Concentration**	**Notes**
N-acetylaspartate (NAA)	Neuronal marker (neuronal density)	~ 8-16 mM	Marker of neuronal integrity
Total creatine (tCr)	Energy buffer	~ 5-10 mM	Marker of energy metabolism

Choline (Cho)	Reflects: membrane synthesis and degradation	~ 1-2.5 mM	Important in cancer diagnostics
Lactate (Lac)	End product of anaerobic glycolysis	< 2 mM	Elevated in pathology
Glutamate (Glu)	Excitatory neurotransmitter	~ 6-12 mM	Important for learning and memory
gamma-aminobutyric acid (GABA)	Inhibitory neurotransmitter	~ 1-2 mM	Balances excitatory activity
Glutathione (GSH)	Antioxidant	~ 2-3 mM	Protects against oxidative stress
Myo-Inositol (mI)	Glial marker, necessary for cell growth	~ 4-8 mM	Important for cell osmoregulation

Another important physical parameter in MRS is relaxation time. The longitudinal relaxation time (T_1) describes the time needed for magnetization to recover along the longitudinal axis (B_0 direction), reaching approximately 63% of its equilibrium value. T_1 depends on the metabolite, tissue type, and field strength, and typically exceeds 1 second (Li et al., 2012). To allow full recovery of longitudinal magnetization, the repetition time (TR) should ideally be more than five times longer than T_1. In practice, however, long TRs increase scan duration, while short TRs reduce the signal-to-noise ratio. A typical TR in single-voxel ^{1}H-MRS experiments is ~1–3 seconds. The transverse relaxation time (T_2) describes the signal decay of transverse magnetization due to microscopic field inhomogeneities and spin–spin interactions. Because the MR signal decays rapidly, it is necessary to select an appropriate echo time (TE) to capture a reliable signal. Spectra acquired with different TEs should not be directly compared, as TE significantly affects signal intensity and metabolite visibility.

A related parameter is the effective relaxation time (T_2*), which describes the signal decay in the free induction decay (FID). In addition to the microscopic processes contributing to T_2, T_2* is also affected by macroscopic magnetic field inhomogeneities. As a result, T_2* is

shorter than the theoretical T_2 and varies depending on the metabolite, tissue composition (gray vs. white matter), field homogeneity, and magnetic field strength (Li et al., 2012; Chavhan et al., 2009). Field inhomogeneities cause faster FID decay (i.e., shorter T_2^* relaxation times), resulting in broader, lower-intensity peaks and distortions in spectral line shapes. To minimize these effects and improve magnetic field homogeneity within the analyzed voxel, a procedure called shimming is applied. It involves passing electrical currents through shim coils to correct for field inhomogeneities. The optimal amount of current in each shim coil is measured and calculated using dedicated algorithms such as FASTMAP (Gruetter, 1993). Detailed recommendations on best practices for achieving magnetic field homogeneity in MRS have been provided in the expert consensus statement by Juchem et al. (2021).

MRS data can be acquired using several types of pulse sequences. One of the most commonly used is PRESS (Point RESolved Spectroscopy), which provides high signal intensity and is relatively robust. However, PRESS sequence, which employs two refocusing pulses, is technically limited in achieving echo times shorter than approximately 30 ms (Klose, 2008; Moonen et al., 1989). An alternative is the STEAM (STimulated Echo Acquisition Mode) sequence, which allows for shorter echo times and more efficient water suppression. However, a stimulated echo sequence provides only half the signal intensity compared to PRESS, and is more sensitive to motion (Moonen et al., 1989; Klose, 2008). A more recent approach is the semi-LASER (Localization by Adiabatic SElective Refocusing) sequence (Deelchand et al., 2021), which is increasingly recommended, particularly at high and ultra-high field strengths, due to its full signal intensity and reduced chemical shift displacement error (CSDE). It has been endorsed by the international MRS consensus group as a preferred localization technique (Wilson et al., 2019). For quantifying metabolites that are present at low concentrations and strongly overlapped with others, such as GABA, spectral editing sequences like MEGA-PRESS are commonly used (Peek et al., 2023).

Several known sources of spectral distortion may affect the quality of MRS data. Field inhomogeneity leads to broader and overlapping peaks but can be mitigated through proper shimming. Subject motion can result in voxel misplacement, frequency or phase shifts, and signal degradation. These artifacts may be corrected during preprocessing by removing motion-corrupted averages and applying frequency and phase alignment

(Andronesi et al., 2021). Another major issue is the CSDE, which causes spatial misalignment of metabolite signals and is especially problematic in single-voxel spectroscopy at high magnetic field strengths (e.g., 7T). This can be addressed by using sequences with high-bandwidth radiofrequency RF pulses (such as semi-LASER) and increased gradient strength**.** Contamination from unsuppressed water or lipids is also a common concern. Water suppression is typically applied during acquisition using methods like VAPOR (Tkáč et al., 2021), and residual water can be removed during preprocessing. These corrections are considered critical in both clinical and research settings. Additional artifacts, such as ghosting (reduced by better field homogeneity) and eddy currents (corrected during preprocessing), should be taken into account during protocol setup and corrected afterward if necessary.

MRS data must be preprocessed to minimize artifacts and prepare for accurate quantification. The following steps are recommended: eddy current correction, removal of motion-corrupted signals, frequency and phase drift correction, spectral alignment, removal of residual water, lipids and spurious echoes, RF coil combination, and signal averaging (Near et al., 2021).

To distinguish individual metabolites in an MRS spectrum, a basis set is used. A basis set is a linear combination of spectral signals from metabolites, each with its own characteristic spectral signature that depends on acquisition parameters such as pulse sequence and echo time (see Figure 2). The number of metabolites included in the basis set is selected based on prior knowledge about which compounds can be reliably detected under the specific acquisition conditions. In addition, macromolecular (MM) signals should be included, especially for short-echo-time sequences, as MM resonances have short T_2 relaxation times, broad frequency spectrum, and can overlap with metabolite signals (Near et al., 2021; Cudalbu et al., 2021). Basis spectra are typically generated using specialized simulation tools or experimental measurements. The acquired MRS data are then fitted to the basis set, and the concentration of each metabolite is estimated as the area under the curve corresponding to its fitted spectral component. The choice and parameterization of the basis set, including the inclusion of macromolecular signals, can significantly influence quantification outcomes, as demonstrated by Hofmann et al. (2002), who showed that metabolite concentrations and fitting errors vary depending on the basis set composition and modeling approach.

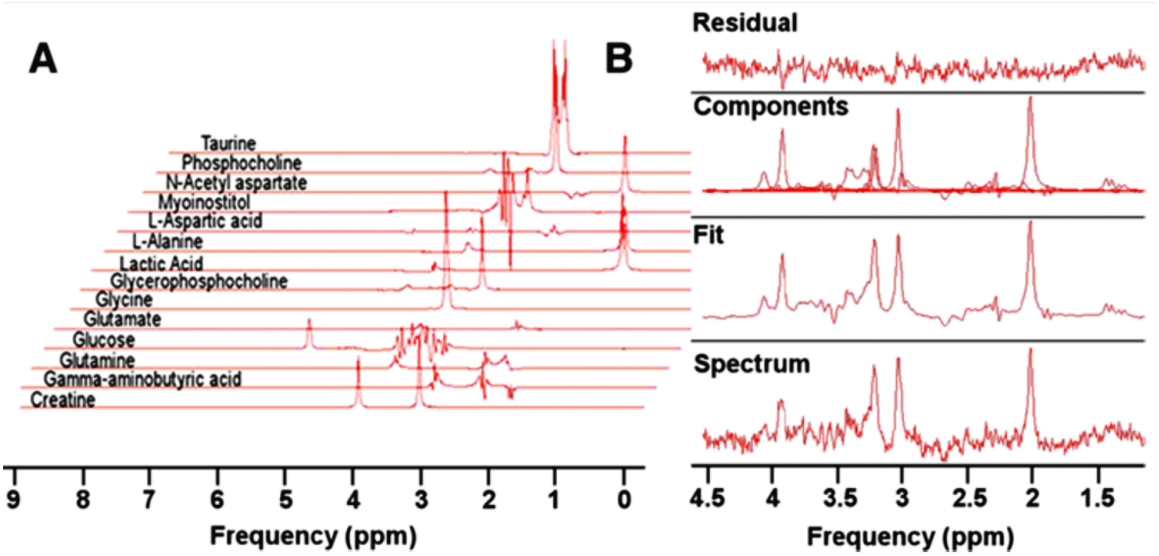

Figure 2. A basis set (left) incorporates chemical shifts, T_2 relaxation times, and line shapes of expected metabolites. The basis set is used to extract a corrected spectrum of the detected components from noisy raw data (lower right). Adapted from Boska et al., 2014, as reproduced on mriquestions.com (under CC BY license).

Metabolite concentrations can be quantified using a water reference or an external sample for absolute quantification, or by using another metabolite as a reference for relative quantification. When using water as a reference, concentrations are typically expressed in molarity (mol/L) or molality (mol/kg), with the latter takes into account variations in water distribution in tissue. One of the main advantages of water referencing is that it can be directly compared between different experiments and settings. Water does not overlap with metabolite peaks, has a concentration approximately 10^4 times higher than that of metabolites, and is affected by the same pulse sequence as the metabolites of interest. Alternatively, another metabolite, such as total creatine (tCr), may be used as a reference, in which case concentrations are reported as ratios. This approach avoids the need for an additional water scan and assumptions related to different relaxation times of water and metabolites (Near et al., 2021).

The quality of the MRS data should be first assessed visually to exclude spectra with major artifacts or those lacking clear metabolite signals. Following visual inspection, the spectra are usually evaluated based on technical parameters. The commonly applied criteria in single-voxel spectroscopy include a linewidth of less than 13 Hz at 3 T, a Cramér-Rao lower bound (CRLB) of 20% or less, and a signal-to-noise ratio (SNR) for N-acetylaspartate (NAA) above 10 for singlet metabolites and above 20 for multiplet signals such as glutamate (Maudsley et al., 2021). Minimum thresholds recommended by

experts to ensure reliable quantification include a CRLB below 50%, an SNR greater than 3 for individual MRS signals to allow for frequency and phase correction, and a linewidth below 19 Hz for 7 T scanners or below 13 Hz for 3 T scanners (Öz et al., 2021).

When comparing magnetic resonance spectroscopy performed at lower and higher magnetic field strengths, such as 3 T versus 7 T (as discussed in Pradhan et al., 2015, and references therein), several advantages of high-field scanners become evident. Magnetic resonance spectroscopy at 7 T benefits from increased signal-to-noise ratio (SNR) and improved spectral resolution (Gruetter et al., 1998; Tkáč et al., 2009). This enhanced spectral quality reduces uncertainty in metabolite quantification (Tkáč et al., 2009; Mekle et al., 2009) and enables the estimation of metabolites present at low concentrations, which at lower fields typically require specialized editing sequences (Mescher et al., 1998). However, high magnetic field strength also introduces technical challenges. These include increased magnetic field inhomogeneities (both B_0 and B_1), which can reduce SNR, as well as longer T_1 and shorter T_2 relaxation times. Additionally, higher field strength leads to increased radiofrequency (RF) power deposition, expressed as the specific absorption rate (SAR), and an elevated CSDE. To address these limitations, partially adiabatic pulse sequences such as semi-LASER are recommended for high-field MRS acquisitions (Öz et al., 2021). Adiabatic pulses provide more uniform flip angles, independent of B_1 field variations, which further improve SNR (Sacolick et al., 2007; Tannus & Garwood, 1997). Furthermore, the use of RF pulses with relatively high bandwidth helps to reduce CSDE. A major advantage of semi-LASER over fully adiabatic sequences like LASER is its reduced SAR, due to a lower number of RF pulses (Garwood & DelaBarre, 2001).

1.1.2. MRS analysis software

As MRS has become more widely adopted in both research and clinical settings, a growing number of software tools for analyzing MRS data have been developed. However, quantification of MRS spectra remains a significant challenge due to their inherent complexity. Spectra are composed of overlapping signals from multiple metabolites, and are often affected by variable lineshapes, unpredictable baselines, and signal contamination from lipids and macromolecules. Artifacts and low signal-to-noise ratios can further complicate interpretation, particularly at lower magnetic field strengths.

Another major obstacle in MRS analysis is the lack of standardized data formats. Each scanner vendor uses its own proprietary file structure: for example, Siemens data may be stored as .dat (Twix), rda, or MRS DICOM files; GE uses .7 (p-files); and Philips typically relies on SPAR/SDAT. On top of that, different analysis software packages support different formats, LCModel requires .RAW files, jMRUI supports .mrui, and FSL-MRS is compatible only with .nii (NIfTI-MRS, Clarke et al. 2022). This diversity in file formats and processing requirements further emphasizes the need for clear and standardized workflows in MRS data analysis. Given the diversity and fragmentation of available tools, it is crucial to understand their capabilities and limitations in order to select appropriate software for specific research purposes.

LCModel is one of the most widely used software packages for ^{1}H-MRS analysis. It automatically fits spectra using the Linear Combination of Model Spectra approach, providing high-precision estimates of metabolite concentrations. The quantification process is highly automated and user-independent, which reduces inter-operator variability and improves reproducibility. Designed to be robust against common MRS data quality issues such as noise, residual water signals, and baseline distortions, the software can still produce suboptimal fits when the acquired spectrum fails to meet minimum signal-to-noise ratio or resolution thresholds. To improve performance in challenging cases, it applies prior constraints on certain metabolite ratios, though these assumptions are not always accurate. In recent years, the developer has made licenses freely available for academic use. Nevertheless, the tool remains closed-source, limiting transparency and preventing full insight into its computational methods. Despite these limitations, LCModel is still regarded as the gold standard for automated ^{1}H-MRS quantification (Provencher, 1993; Provencher, 2001).

Tarquin is another popular software package for analyzing ^{1}H-MRS data. It is a free, open-source program written in C++, originally developed to support both in vivo MRS and ex vivo high-resolution magic angle spinning (HR-MAS) spectroscopy, particularly for the study of brain tumors and other pathologies (Wilson et al., 2011). It features an intuitive graphical user interface (GUI), which allows users to easily perform basic MRS analyses. However, Tarquin does not currently support the Siemens .twix file format, and the software has not been actively developed in recent years, the last version was released in 2018 (https://tarquin.sourceforge.net/).

jMRUI is a software package dedicated to the analysis of MRS and MRSI data, developed in Java. It features a user-friendly graphical interface, which facilitates access to the program's advanced processing and quantification capabilities (Mocioiu et al., 2015; Stefan et al., 2009). The software is freely available for non-commercial and academic use upon registration and acceptance of a license agreement (http://www.jmrui.eu).

Osprey is a free, open-source, MATLAB-based software package for MRS data analysis. It was developed to improve standardization, transparency, and accessibility in MRS data processing pipelines. Osprey supports a wide range of editing and non-editing sequences, including MEGA-PRESS, making it suitable for analyzing GABA and other edited metabolites (Oeltzschner et al., 2020). The software integrates multiple processing steps, from raw data conversion to quantification and quality control, in a single pipeline and it is well-suited for both single-subject and group-level analyses.

Gannet is a free, open-source MATLAB-based software tool specifically designed for analyzing edited single-voxel ^{1}H-MRS data, with a primary focus on GABA quantification. It offers a largely automated analysis pipeline, requiring minimal user interaction, and supports the MEGA-PRESS sequence commonly used in edited MRS acquisitions (Edden et al., 2014). Gannet has become a widely used tool in both clinical and research settings for its simplicity and reliability in GABA+ signal quantification.

FID-A (FID Appliance) is an open-source MATLAB-based software package designed for simulating and processing MRS data (Simpson et al., 2017). The toolbox is particularly useful for converting raw MRS data from vendor-specific formats into formats compatible with other commonly used processing tools. In addition, FID-A enables the simulation of custom basis sets, which can be integrated into quantification pipelines such as LCModel or Tarquin. Its flexibility and open architecture make it a valuable resource for advanced MRS research and method development.

Suspect is a free, open-source Python-based toolkit for processing MRS data, particularly aimed at preparing datasets for quantification using LCModel and other tools. It offers a complete preprocessing pipeline based on the expert consensus recommendations for MRS data handling (Near et al., 2021). The pipeline is implemented in an interactive Jupyter Notebook environment, allowing for easy, real-time modifications and visualization. Suspect is especially useful for researchers who prefer flexible, scriptable

workflows and want to integrate MRS processing with modern Python-based analysis ecosystems.

Vespa (Versatile Simulation, Pulses, and Analysis) is a Python-based, open-source toolkit designed for spectral simulation, interactive data processing, and RF pulse design in magnetic resonance spectroscopy (Soher et al., 2023). The software enables users to create synthetic MRS datasets, prototype custom acquisition schemes, and simulate basis sets compatible with quantification tools such as LCModel. Vespa also includes a powerful RF pulse design module, making it a useful platform for developers working on novel MRS sequences and methods.

Finally, FSL-MRS, is a free, open-source, Python-based toolbox for end-to-end analysis of MRS data (Clarke et al., 2021). It provides a modular and fully integrated pipeline, including data conversion (via spec2nii), preprocessing, spectral simulation, model fitting, quantification, and advanced visualization. Released in June 2020, FSL-MRS is a highly active and rapidly evolving project, with over 60 version updates in March 2025. Its transparent architecture and full compatibility with the NIfTI-MRS standard make it particularly well-suited for reproducible MRS research. Importantly, FSL-MRS is currently the only MRS software package equipped with a dedicated dynamic fitting module, specifically designed for fMRS data analysis (Clarke et al., 2024). This makes it a unique and powerful tool for studying neurochemical dynamics during cognitive or sensorimotor tasks.

In addition to the major software packages described above, there are several other tools developed for MRS data processing and analysis. These include:

- CloudBrain-MRS: a cloud-based computation platform that integrates LCModel and advanced deep learning algorithms for denoising and quantification (Chen et al., 2024),
- MRSCloud: a web-based application for simulating basis sets (Hui et al, 2022),
- MRS DeIdentification Tools – designed to remove protected health information (PHI) from MRS data (https://github.com/schorschinho/MRSDeIdentificationTools),
- INSPECTOR: a versatile platform for processing, analyzing, and visualizing MRS data (Gajdošík et al., 2021),

- MARSS: a tool for fast and high-precision basis set simulation (Landheer et al., 2021),
- MRspa: a software for post-processing and quantification of MRS data (Deelchand D. MRspa: Magnetic Resonance signal processing and analysis. Available at: https://www.cmrr.umn.edu/downloads/mrspa/).

The sheer number and diversity of MRS-related tools reflects the growing interest in MRS. However, it also highlights a lack of standardization in the field: most of these programs are not widely adopted, and few researchers use multiple tools concurrently. This fragmentation may limit the reproducibility of analyses and the community's ability to detect methodological errors. Furthermore, many of these tools were developed as part of individual research projects, and are non-commercial, which may impact long-term support and documentation

1.1.3. Functional magnetic resonance spectroscopy

Functional magnetic resonance spectroscopy (fMRS) is an innovative neuroimaging technique that enables the measurement of metabolite concentrations during brain activation in response to various types of stimulation. In contrast to conventional MRS, which typically captures resting-state metabolite levels, fMRS allows researchers to investigate dynamic, time-resolved neurochemical responses that occur during cognitive, sensory, or motor tasks. This ability to capture the temporal dimension of neurochemical changes offers unique insights into the biological mechanisms underlying brain function (Stanley & Raz, 2018). In particular, fMRS has proven especially valuable in detecting task-related modulations in glutamate, a key excitatory neurotransmitter involved in synaptic plasticity, learning, and cognitive processing.

Although fMRS is based on the same physical principles, pulse sequences, and acquisition parameters as standard MRS, it differs in its experimental paradigm. Participants typically perform simple tasks, visual, motor, or cognitive, while remaining as still as possible to avoid motion-related artifacts. Data are acquired in alternating periods of task and rest, with rest intervals assumed to reflect baseline metabolite levels. Importantly, individual MRS signals (transients) are analyzed separately rather than averaged across the entire acquisition. Depending on the study design, these signals can be grouped by condition (e.g., task vs. rest) or modeled individually in a time-resolved manner. Most fMRS studies

target glutamate, alone or combined with glutamine (Glx), as well as GABA, the main inhibitory neurotransmitter.

Despite the promising potential of fMRS, several methodological challenges must be addressed to ensure accurate and interpretable results. One of the main challenges in fMRS is contamination of the spectra by the blood oxygenation level–dependent (BOLD) effect. BOLD-related changes in local magnetic fields can affect spectral line shapes. These effects typically manifest as line narrowing and associated amplitude changes, which may reduce the accuracy of metabolite quantification (Stanley & Raz, 2018; Mangia et al., 2006). Several strategies have been proposed to reduce these effects. An early method was line broadening correction, where activation spectra are artificially broadened to match the baseline spectrum and minimize fitting errors (Mangia et al., 2007; Schaller et al., 2014). Another approach is the simultaneous acquisition of BOLD and fMRS, which allows direct modelling of the BOLD signal alongside neurochemical data, helping to separate vascular from metabolic effects (Ip et al., 2017; Ip et al., 2019). More recently, model-based approaches have been developed, using general linear models with the hemodynamic response function as a regressor during spectral fitting. This enables direct estimation of metabolite changes while accounting for BOLD effects (Clarke et al., 2024). However, the BOLD response in fMRS is still less well understood than in functional magnetic resonance imaging (fMRI). Its influence on quantification reliability and interpretation remains uncertain, especially in regions with low signal-to-noise ratio or in event-related designs (Liu et al., 2025). This uncertainty poses a major challenge for standardization and broad application of fMRS in cognitive neuroscience.

Another critical limitation of fMRS is its relatively low signal-to-noise ratio (SNR) and the subtle magnitude of metabolite concentration changes, which together result in reduced sensitivity (Liu et al., 2025). To overcome this, fMRS studies often require larger sample sizes, sometimes several dozen participants, to achieve sufficient statistical power. Additionally, longer acquisition times help improve SNR by increasing the number of MRS transients collected. However, prolonged scanning increases the risk of motion artifacts and may reduce participants' ability to maintain attention and task engagement. Another challenge associated with longer acquisitions is frequency drift, which can compromise spectral quality over time. fMRS is highly sensitive to magnetic field inhomogeneities, and even small fluctuations, arising from frequency drift or physiological factors such as

breathing, can affect quantification accuracy. Since individual MRS signals (transients) are analyzed separately, these subtle shifts in magnetic field can propagate into the final results and lead to misinterpretations.

Although several mechanisms have been proposed to explain metabolite changes observed with fMRS, their exact biological origins remain unclear (Pasanta et al., 2023; Buxton, 2009). Proposed explanations include increased energy metabolism via the tricarboxylic acid (TCA) cycle (Dienel, 2012; Magistretti & Allaman, 2015), neurotransmitter release into the synaptic cleft during neuronal activation (Buxton, 2009), astrocyte-mediated glutamate–glutamine cycling (Sibson et al., 1998; Rothman et al., 2003), and longer-term neuroplastic processes involving brain-derived neurotrophic factor (BDNF) signaling (Gonçalves-Ribeiro et al., 2019; Valtcheva & Venance, 2019). Additionally, GABA levels may reflect a shift in excitation/inhibition balance (Just et al., 2013; Lynn et al., 2018) or general inhibitory tone rather than phasic neurotransmission (Rae, 2014; Peek et al., 2020). Since fMRS captures the total metabolite pool within the voxel, it is currently not possible to disentangle synaptic, metabolic, and vesicular contributions to the measured signal, warranting caution in interpretation (Puts & Edden, 2012; Takado et al., 2021). Furthermore, these mechanisms are likely to operate on different timescales and may contribute to the measured signal at varying points following task onset.

A key challenge in interpreting fMRS findings is the limited understanding of how metabolite concentrations, particularly glutamate, change over time following stimulation. This uncertainty complicates experimental design, especially the timing of signal acquisition relative to task onset. In an effort to address this, Mullins (2018) proposed a theoretical glutamate response function (GRF) based on a meta-analysis of existing fMRS studies (see Figure 3). Although the GRF serves as a useful conceptual tool, the actual dynamics of glutamate in the human brain remain unknown, due in part to the unresolved biological mechanisms (hypothetically related to neurotransmission).

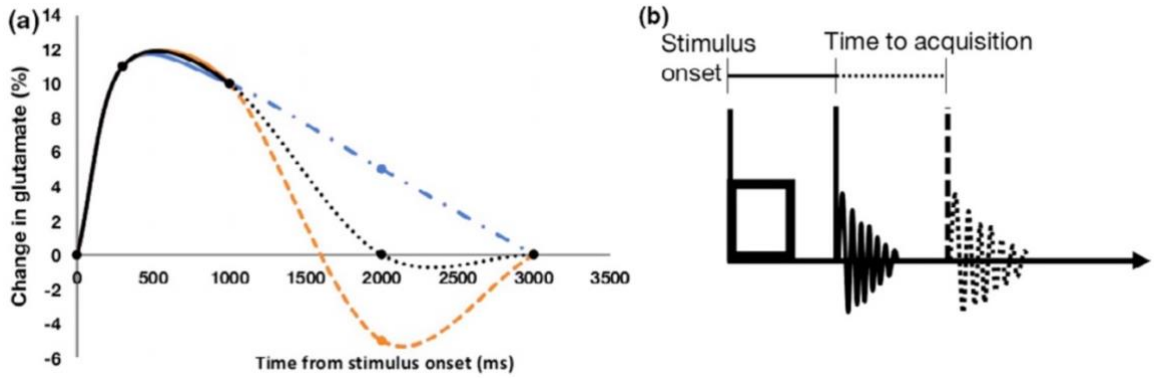

Figure 3. Hypothesized Glutamatergic Response Function (GRF). Adapted from Mullins (2018), under CC BY license.

An important methodological decision in fMRS studies concerns the choice between block and event-related designs (Apšvalka et al., 2015; Koolschijn et al., 2023). In block designs, spectra are acquired in alternating periods of stimulation and rest, typically lasting several minutes each. Signals from each condition are averaged separately, and the resulting spectra are compared. While this approach improves signal-to-noise ratio, it offers limited temporal resolution. In contrast, event-related designs involve the acquisition of individual transients for each stimulus, with conditions presented in a randomized or intermixed fashion. This allows for higher temporal resolution and finer modeling of metabolite dynamics, but substantially increases analytical complexity. There is currently no consensus on which fMRS design is superior for capturing task-related neurochemical dynamics. A meta-analysis by Mullins (2018) suggested that event-related designs reveal greater glutamate responses to stimulation, likely due to better temporal resolution. Similarly, Pasanta et al. (2023), in a comprehensive review of 49 studies, reported larger effect sizes for event-related paradigms. However, they also found that block designs produced more consistent and reproducible results, likely due to higher signal-to-noise ratios. These findings highlight a trade-off between sensitivity and reliability, and underscore the importance of aligning the paradigm choice with the research question and analytical capabilities.

Functional magnetic resonance spectroscopy (fMRS) has been used to study many brain regions, stimulation types, and metabolites, with mixed results. In the visual cortex, early studies reported rapid lactate increases during flashing light stimulation (Prichard et al., 1991). Later work in V1 found small but significant glutamate (or Glx) increases during visual checkerboard tasks (Ip et al., 2017; Yakovlev et al., 2022), and decreases in GABA

were reported in the dark, without visual input (Kurcyus et al., 2018). Motor tasks did not show significant glutamate increases in the motor cortex (Kolasinski et al., 2019), whereas sustained finger tapping was associated with combined lactate and glutamate rises (Schaller et al., 2014). Pain-related tasks in the cingulate cortex have consistently shown glutamate increases during thermal or pressure pain (Archibald et al., 2020; Jelen et al., 2021). Cognitive and imagery tasks have produced region-specific effects, such as higher Glx in the medial prefrontal cortex of competitive swimmers during motor imagery (Huang et al., 2015) or glutamate increases in the dorsolateral prefrontal cortex during high working-memory load (Woodcock et al., 2018). Social processing studies have not reported higher glutamate in the superior temporal sulcus when viewing dynamic social stimuli compared to non-social stimuli, with observed changes instead related to visual processing in V1 (Pasanta et al., 2024). Overall, this research shows that fMRS can detect task-related neurochemical changes, but results vary, and methodological differences remain a major challenge.

Among the many cognitive domains that could be explored using fMRS, reading is particularly well suited for investigation. Reading engages a well-characterized network of brain regions that could serve as robust regions of interest for neurochemical measurement. By targeting these regions, fMRS offers the potential to reveal task-related metabolite changes that may underlie reading skill and its impairments. However, to date, no studies have directly investigated this topic.

1.2. Dyslexia

Dyslexia is a specific learning disorder that causes difficulties in the acquisition of reading skills. These difficulties cannot be explained by lower intelligence, vision impairment, poor access to education, or inappropriate teaching methods. Children with dyslexia often experience lower self-esteem, anxiety, depression, and a lack of confidence in their reading abilities (Wilmot et al., 2023).

In every language, 5-12% of children struggle with slow and inaccurate reading that does not improve over time. Interestingly, many studies suggest that developmental dyslexia occurs more frequently in males than in females (Arnett et al., 2017). Additionally, the risk of dyslexia increases to 45% among children who have a first-degree relative with reading difficulties (Snowling & Melby-Lervåg, 2016). This points to a genetic factor, with some

studies proposing specific genes involved in dyslexia (Fisher & Francks, 2006; Bieder et al., 2020). Despite decades of research on reading disorders, the underlying causes of dyslexia remain unknown, and many theories have been proposed to explain the mechanisms behind developmental dyslexia.

One theory suggests that a deficit in phonological processing underlies dyslexia (Schwarz et al., 2024; Ramus et al., 2013; Snowling, 2001). This theory posits that dyslexia arises from difficulties in processing the sounds of language (phonemes) and associating them with letters (graphemes). Individuals with dyslexia may struggle with recognizing and manipulating these sounds, which are essential for reading and spelling. Phonological processing is critical not only for decoding words but also for encoding them, which impacts reading fluency, writing, and spelling accuracy. These difficulties often manifest as trouble with tasks like rhyme detection, phoneme segmentation, and non-word repetition. Additionally, while phonological processing deficits are an indicator of dyslexia, the severity and specific nature of these deficits can vary, with some individuals exhibiting better compensatory strategies than others, such as relying more on semantic or visual memory. Based on the hypothesis that phonological processing difficulties lie at the core of developmental dyslexia, an extended theory known as the double-deficit hypothesis has been proposed. This hypothesis suggests that dyslexia may result from deficits in both phonological awareness and rapid automatized naming (RAN), the latter referring to the ability to name visual stimuli aloud as quickly as possible (Wolf & Bowers, 1999). Supporting this view, a study conducted on Polish-speaking children found that 51% of those with dyslexia exhibited a phonological deficit, while 26% showed a RAN deficit. Notably, both deficits coexisted in 14% of the dyslexic sample (Dębska et al., 2022).

Another theory regarding the origin of dyslexia is the visual processing deficit hypothesis, which proposes impairments in visual processing, particularly in the dorsal visual stream that is responsible for spatial attention and the sequential processing of letters. Individuals with dyslexia may have difficulties perceiving and accurately recognizing letters, especially when tasks require rapid or sequential visual processing. This visual attention deficit can contribute to difficulties with reading, including issues with phoneme-grapheme correspondence (Vidyasagar & Pammer, 2010).

Recently, the neural noise hypothesis has been proposed, suggesting that dyslexia might be caused by neuronal hyperexcitability, which leads to difficulties in distinguishing relevant

signals from background noise (see Figure 4). This theory posits that increased concentrations of glutamate (Glu), the primary excitatory neurotransmitter, contribute to this hyperexcitability, impairing the precision of neural firing and the synchronization of neural networks, particularly in the left superior temporal cortex, a region critical for phonological processing. As a result, individuals with dyslexia may struggle with the accurate processing of auditory and visual stimuli, leading to challenges in phonological processing and the integration of visual symbols with their corresponding speech sounds. The hypothesis also suggests that genetic risk factors, such as mutations in the DCDC2 and KIAA0319 genes, could influence neural excitability and contribute to the development of neural noise. Understanding the role of neural noise in dyslexia offers potential avenues for intervention, including brain stimulation techniques (e.g., tDCS, TMS) or pharmacological treatments that modulate neural excitability and restore the balance between excitatory and inhibitory signals (Hancock et al., 2017).

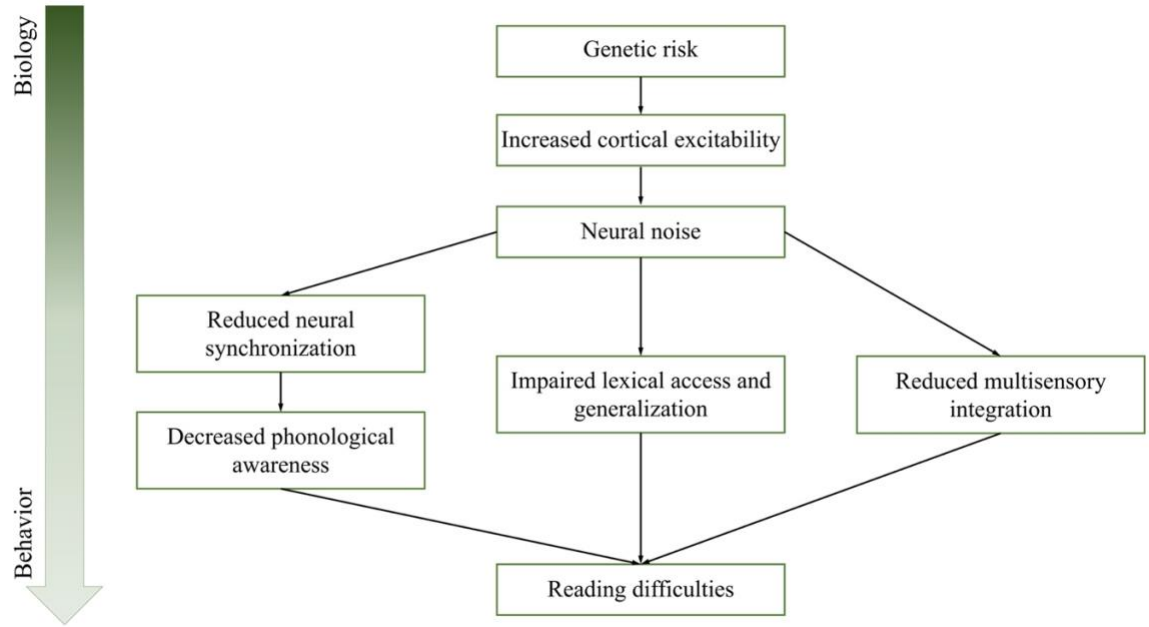

Figure 4. Simplified schematic of the neural noise hypothesis in dyslexia (based on Figure 1 from Hancock et al., 2017).

1.2.1. The neural basis of reading and dyslexia

Although the reading network encompasses multiple left-hemispheric brain regions (Dehaene, 2009), the most critical areas include the visual word form area (VWFA) for orthographic processing, the superior temporal gyrus/sulcus (STG/STS) for phonological processing, and the inferior frontal gyrus (IFG) for higher-order language integration.

The Visual Word Form Area (VWFA), a region located in the left fusiform gyrus of the ventral occipito-temporal cortex, is critically involved in the recognition of written words and is considered a key component of the neural circuitry underlying fluent reading. This region exhibits stronger responses to visually presented words than to other types of visual stimuli matched for complexity. Functional imaging studies have shown that the VWFA is more strongly activated by familiar scripts compared to line drawings, digit strings, or unfamiliar characters, indicating its specialized role in processing orthographic information (Dehaene & Cohen, 2011). In children with dyslexia, reduced VWFA activation during reading is a consistent finding across many orthographies (see Chyl et al., 2021 for review). Functional MRI data from a large cohort of school-aged children revealed significantly lower word-related activation in this area also among poor readers (Brem et al., 2020). Furthermore, the degree of VWFA hypoactivation was proportional to the severity of reading difficulties, suggesting a strong link between visual word form processing and impaired reading development (Brem et al., 2020). In addition to its role in visual word recognition, the VWFA demonstrates functional and structural connectivity with both language-related and attentional brain networks (Chen et al., 2019). Analysis of large-scale neuroimaging data has revealed that connectivity between the VWFA and the lateral temporal language network predicts linguistic skills, whereas its connectivity with the dorsal fronto-parietal attention system is associated with visuo-spatial attention. This pattern of distinct and functionally meaningful connections supports a multiplex model in which the VWFA integrates linguistic and attentional processes, rather than operating solely within the reading network (Chen et al., 2019).

The left superior temporal sulcus (STS), located in the posterior region of the superior temporal lobe between the superior and middle temporal gyri, plays a key role in language processing as well as cross-modal integration, particularly in mapping auditory and visual linguistic input. It responds to both auditory and visual stimuli, reflecting its function as a multisensory hub (van Atteveldt et al., 2004). This region is considered essential for the

formation of robust letter–speech sound (LSS) integration, a process fundamental to early reading acquisition. In typical readers, this process is characterized by robust congruency effects, i.e., differential activation in response to congruent compared to incongruent LSS pairs. Effective LSS integration is linked to bilateral activity in the heteromodal STS and surrounding auditory regions, including the planum temporale (Richlan et al., 2019). Findings from longitudinal studies support the involvement of STS in reading development. For instance, a study of Polish-speaking children aged 7 to 9 revealed increased activation in the left superior temporal cortex and bilateral inferior frontal cortex for multisensory over unisensory LSS stimuli with time and reading experience (Beck et al., 2024). In individuals with dyslexia, reduced activation of the superior temporal gyrus (STG) and STS has been observed during audiovisual tasks, suggesting impaired responsiveness to multisensory input (Ye et al., 2017). Additionally, abnormalities in gray matter distribution within the temporal voice area, a subregion of the STS, have been reported in adults with dyslexia (Dole et al., 2013). These findings are consistent with broader evidence of reduced gray matter volume in the left STS and right STG in dyslexia across different alphabetic orthographies (Richlan et al., 2020). A recent meta-analysis including both alphabetic and logographic orthographies has confirmed the left STG/STS as a region consistently showing hypoactivation and reduced gray matter volume in dyslexia across many languages (Yan et al., 2021).

Finally, the left inferior frontal gyrus (IFG), particularly within Broca's area, plays a foundational role in reading by mediating phonological decoding, lexical retrieval, and the integration of orthographic and semantic information (Heim et al., 2005). Similar to the left superior temporal gyrus (STG), it is part of the language network and becomes involved in reading following literacy acquisition (Chyl et al., 2018). The exact nature of left IFG dysfunction in developmental dyslexia remains controversial. While numerous fMRI and meta-analytic studies have documented reduced activation of the left IFG in individuals with developmental dyslexia (e.g., Martin et al., 2016), other investigations have reported increased activation in this region (e.g., Georgiewa et al., 2002; Hoeft et al., 2007). Such discrepancies may reflect differences in task type and difficulty, orthographic transparency (Martin et al., 2016; Yan et al., 2021). Moreover, structural neuroimaging studies have reported reduced gray matter volume in the left IFG (e.g., Jednoróg et al., 2014), although this finding is less consistently observed than reductions in the left STG/STS in meta-analyses.

1.2.2. Sex differences in dyslexia

Dyslexia is diagnosed more often in males than in females, with a ratio of about 3:1 (Yang, 2022) or even 5:1 (Arnett et al., 2017). Historically, more male subjects were included in studies (65% in children and 95% in adults), and the assumption that females have the same dyslexic profile could be incorrect (Krafnick & Evans, 2019). Nowadays, when study samples are balanced with regard to sex, the framework for interpretation is still often based on theories constructed for cohorts with a male predominance (Krafnick & Evans, 2019).

Previous studies have found significantly greater variability and lower mean performance in reading skills for males, which leads to their overrepresentation in the low-performance tail of the reading distribution (Arnett et al., 2017). It has also been shown that males have slower processing speeds and worse inhibitory control but perform better in verbal reasoning tasks (Arnett et al., 2017). However, results in phonemic awareness or working memory did not differ between the sexes (Arnett et al., 2017). On the other hand, Krafnick and Evans (2019) point out that females with dyslexia tend to have higher IQ scores, perform better in working memory tasks, and are more proficient in orthographic coding. They also observed sex differences in visual prediction tasks, where males with dyslexia focused on the current target, while females were able to predict subsequent targets more quickly (Krafnick & Evans, 2019). To date, only few MRI study have included sex as a variable of interest when examining dyslexia effects. For example, Evans and colleagues (2014) reported that dyslexic males exhibited reduced gray matter volume in the left temporo-parietal cortex, whereas in females, differences were found outside the canonical reading network, primarily in the right hemisphere.

1.2.3. MRS studies of dyslexia

Several MRS studies have examined brain metabolite concentrations in individuals with dyslexia. The first study employing proton magnetic resonance spectroscopy (^{1}H-MRS) was conducted on 29 adult males (14 with dyslexia and 15 typical readers), with single-voxel MRS data acquired bilaterally from the temporo-parietal cortex and the cerebellum (Rae et al., 1998). A significantly decreased choline-to-N-acetylaspartate (Cho/NAA) ratio was observed in the left temporo-parietal cortex and the right cerebellum in the dyslexic group. Additionally, a lateralized biochemical difference was found in both brain regions among dyslexic participants, which was not present in controls. A year later,

an investigation used magnetic resonance spectroscopic imaging (MRSI) to examine brain lactate metabolism during language-related and non-language cognitive tasks (Richards et al., 1999). The study included six boys with dyslexia and seven typically developing boys. MRSI data were collected while participants performed three phonological tasks and one non-language task. Results showed that, during phonological tasks, children with dyslexia exhibited a greater number of voxels with elevated lactate levels in the left anterior region of the brain compared to controls. No group differences were observed during the non-language task. Lactate metabolism was evaluated using the lactate-to-N-acetylaspartate (Lac/NAA) ratio. Another study investigated cerebellar metabolite concentrations in six adult males with dyslexia and six matched controls (Laycock et al., 2008). A decreased NAA/Cho ratio in the right cerebellar hemisphere was reported, along with an increased Cho/Cr ratio in the left hemisphere, suggesting hemispheric differences in cerebellar metabolism in dyslexia.

Further research used ^{1}H-MRS to examine relationships between brain metabolite levels and reading abilities in 31 adults (17 females, 14 males) with varying reading proficiency (Bruno et al., 2013). A negative correlation was found between phonological decoding and the Cho/Cr ratio in the left angular gyrus, suggesting that higher choline levels in this region may be associated with less efficient phonological processing. Additionally, higher NAA/Cr levels and greater general cognitive ability were reported in males, although the results may reflect sampling bias due to the small sample size. A larger sample of 75 children (47 boys, 28 girls) was assessed to investigate neurometabolite concentrations during the developmental period of reading acquisition (Pugh et al., 2014). Results indicated that higher Cho/Cr and Glu/Cr levels in the visual cortex were negatively correlated with reading ability and phonological processing. Additional evidence came from a study involving children aged 3.0 to 5.4 years (Lebel et al., 2016). Spectra from the anterior cingulate gyrus were analyzed in 56 participants (32 boys, 24 girls), revealing positive correlations between phonological processing and glutamate, creatine, and inositol concentrations. In the left angular gyrus (45 children: 29 boys, 16 girls), rapid automatized naming showed trend-level negative correlations with glutamine and choline. Children with lower levels of these metabolites tended to perform better on the naming task, indicating better language skills. In this study, metabolite concentrations were expressed in absolute units (mmol/kg), with water used as the reference. No significant sex differences were found in language task performance or metabolite levels. Another study investigated

GABA concentrations bilaterally in the inferior frontal gyrus in 28 adults (Nakai & Okanoya, 2016). No hemispheric differences in GABA+/Cr levels were found. However, a negative correlation between GABA+/Cr levels in the left inferior frontal gyrus and verbal fluency scores was reported. An additional investigation comes from an analysis of 70 children (44 boys, 26 girls), with MRS data acquired from the midline occipital cortex (Del Tufo et al., 2018). Participants taking part in this study were a subset of the sample described in Pugh et al. (2014). Elevated glutamate and choline levels were associated with poorer reading performance, while lower GABA and higher N-acetylaspartate (NAA) concentrations were related to faster reaction times in a reading-related task. All metabolite levels were reported as ratios relative to creatine (Cr), which served as the internal reference. In the same year, another study reported a negative correlation between choline and myo-inositol concentrations and processing speed in girls with dyslexia, a pattern not observed in boys (Horowitz-Kraus et al., 2018). The sample included 24 children with dyslexia (8 girls) and 30 typical readers (17 girls). MRS data were acquired from the perigenual anterior cingulate cortex (ACC), a region associated with executive control and attention. A notable study explored neurometabolite concentrations in a Polish-speaking sample of 36 adults (18 with dyslexia, 18 controls) and 52 children (26 with dyslexia, 26 controls) (Kossowski et al., 2019). MRS data were collected from the left temporo-parietal and occipital cortex. In both age groups, individuals with dyslexia showed lower total NAA (tNAA) concentrations in the occipital cortex, indicating potential reductions in neuronal integrity or density. The study also demonstrated age-related differences in metabolite levels: adults had higher choline and creatine in both regions, higher tNAA in the temporo-parietal cortex, and lower glutamate in the occipital cortex. Importantly, creatine, often used as a stable reference metabolite, was shown to vary with age. The most recent investigation explored neurometabolite concentrations in the anterior cingulate cortex in 21 children with dyslexia (14 boys, 7 girls) and 31 typical readers (14 boys, 17 girls) (Cecil et al., 2021). Better word reading performance in the dyslexic group was associated with lower concentrations of Glx, Glu, Cr, and NAA. As in the study by Horowitz-Kraus, metabolite concentrations were reported in millimolar units and scaled to water, with corrections applied for tissue composition and relaxation times, except for Glu and Glx, which were not adjusted for metabolite-specific relaxation.

In addition, phosphorus-31 magnetic resonance spectroscopy (^{31}P MRS) was employed to investigate cerebral phospholipid metabolism in adults with dyslexia (Richardson et al.,

1997). The study included 12 individuals with dyslexia (7 males, 5 females) and 10 controls (5 males, 5 females). Spectra were acquired from ten brain voxels per participant, though specific anatomical locations were not clearly reported. The dyslexic group showed a significantly elevated phosphomonoester (PME) peak, indicating potential abnormalities in membrane phospholipid metabolism.

Overall, findings from MRS studies in dyslexia remain inconsistent, likely due to methodological variability and small, diverse samples. A lack of standardization in regions of interest and reference methods further limits cross-study comparability.

1.2.4. Evidence related to the neural noise hypothesis

The recently proposed neural noise hypothesis of dyslexia (Hancock et al., 2017) suggests that poor reading skills are associated with cortical hyperexcitability. According to this theory, individuals with dyslexia should exhibit elevated concentrations of the primary excitatory neurotransmitter, glutamate (Glu), and reduced levels of the main inhibitory neurotransmitter, gamma-aminobutyric acid (GABA). The resulting imbalance between excitatory and inhibitory signaling is believed to disrupt the temporal precision and synchronization of neuronal firing, leading to increased background neural "noise" that interferes with precise signal transmission.

Empirical findings related to this hypothesis remain mixed. For example, one study reported that higher glutamate concentrations in the midline occipital cortex were negatively correlated with reading and phonological abilities (Pugh et al., 2014). The authors were also among the first to suggest that cortical hyperexcitability, together with disrupted white matter organization, may underlie dyslexia. A follow-up analysis on a largely overlapping sample (70 of the 75 participants) extended the findings by showing that elevated glutamate in the same occipital region was again linked to poorer reading performance, while lower GABA concentrations were associated with faster response times in a cross-modal matching task (Del Tufo et al., 2018). Additional support comes from research showing that reduced Glx and Glu concentrations were observed in children with dyslexia who performed better on a word reading task following a reading intervention targeting executive function (Cecil et al., 2021).

In contrast, other findings do not align with the predictions of the neural noise hypothesis. One investigation found that higher glutamate concentrations in the anterior cingulate cortex were actually associated with better phonological skills (Lebel et al., 2016). Similarly, increased GABA+/Cr levels in the left inferior frontal gyrus were related to poorer verbal fluency (Nakai & Okanoya, 2016), a common challenge for individuals with dyslexia.

There are also studies that found no evidence for a relationship between Glu or GABA concentrations and reading skills. For instance, an investigation focused on the anterior cingulate cortex reported no significant correlations between metabolite levels and either reading ability or cognitive task performance (Horowitz-Kraus et al., 2018). Another study that examined the left temporo-parietal and occipital cortices found no group differences in Glu, Glx, or GABA levels between dyslexic and typically developing individuals (Kossowski et al., 2019).

1.3. Summary

Among the many cognitive domains that could be investigated with functional magnetic resonance spectroscopy (fMRS), reading offers a particularly compelling target. Reading engages a well-characterized left-hemispheric network of brain regions—including the visual word form area (VWFA), superior temporal sulcus (STS), and inferior frontal gyrus (IFG)—that provide robust regions of interest for probing task-related neurochemical changes.

To date, however, no studies have applied fMRS to the reading process, despite the potential of this approach to reveal dynamic glutamate responses that may underlie both typical reading and its impairments. Prior work has shown that salient stimuli are required to evoke reliable glutamate responses (Ip et al., 2019), with detection enhanced at ultra-high-field (7T) relative to conventional 3T scanners (Terpstra et al., 2016). Visual word forms, in particular, constitute highly salient stimuli that capture attention from early childhood (Zhao et al., 2018) and consistently activate the reading network (Dehaene & Cohen, 2011). This makes them ideally suited for testing whether fMRS can detect task-modulated neurochemical changes during reading.

Importantly, it remains unknown whether such modulation relates to individual differences in reading skill, or whether fMRS can distinguish between typical and dyslexic readers. While resting-state MRS studies of dyslexia have produced mixed evidence for elevated glutamate levels (e.g., Pugh et al., 2014; Kossowski et al., 2019), it is plausible that neuronal hyperexcitability, if present, might be more detectable during reading-related stimulation and confined to regions of the reading network rather than more general cortical areas.

2. Original Study

2.1. Aims and hypotheses

In this thesis, I address five research aims that explore the use of fMRS to study glutamate concentration changes during reading in typical and dyslexic readers, while also assessing the regional specificity, temporal dynamics, and data quality across ultra-high-field and conventional scanners.

The first goal was to examine whether glutamate concentration changes in response to reading-related visual stimuli within language-sensitive brain regions, specifically the superior temporal sulcus (STS) and the visual word form area (VWFA). To the best of my knowledge no previous single-voxel fMRS study has used reading-related stimuli with voxels placed in individually defined regions directly involved in reading. It was expected that words would evoke a stronger glutamate response than false font strings, particularly in the VWFA (Dehaene & Cohen, 2011), and that post-stimulation glutamate levels would be higher than during rest.

The second aim was to test the neural noise hypothesis, which predicts that individuals with dyslexia exhibit higher glutamate levels in the STS compared to typical readers (Hancock et al.). In this case, a significant between-group difference in glutamate concentration was expected in this region, with the dyslexic group showing elevated levels, especially during reading word forms.

The third aim was to determine whether the effects described above are specific to brain regions associated with reading (VWFA and STS) or whether they can also be observed in the medial prefrontal cortex (mPFC), which is not directly involved in reading (Richlan et al., 2009). This comparison would allow assessment of the regional specificity of the observed effects.

A further aim was to investigate the temporal dynamics of glutamate concentration changes following stimulation. Despite increasing interest in fMRS, the exact shape of the glutamate response function remains unknown. The theoretical function proposed by Mullins (2018) presents three potential patterns. However, they still need to be empirically confirmed. In the present study, glutamate concentration was measured at four time points

after stimulus onset: 500 ms, 1000 ms, 3000 ms, and 4500 ms. The first two time points were hypothesized to capture the peak and sustained elevation of glutamate, while the latter two were expected to reflect its return to baseline levels.

Finally, the study aimed to compare data quality obtained on two different magnetic resonance systems: a 7 Tesla DISCOVERY 950 MR System (GE, Ultra-High Field Magnetic Resonance Lab, ECO-TECH Complex, Lublin, Poland) and a 3 Tesla Trio system (Siemens, Laboratory of Brain Imaging, Nencki Institute of Experimental Biology, Warsaw, Poland). Ultra-high-field scanners are expected to provide higher signal-to-noise ratio and better spectral resolution, enabling more accurate metabolite quantification. However, they also pose challenges, such as magnetic field inhomogeneity and potential signal loss due to shorter T_2 and longer T_1 relaxation times (Pradhan et al., 2015). Assessing whether the technical advantages outweigh the practical limitations was therefore an important part of this work.

To address these aims, three experiments were conducted. The first experiment used 7T data to measure glutamate responses to words, false font strings, and rest, and additionally glutamate concentration changes across the four defined time points. The second experiment followed the same design at 3T. The third experiment compared data quality between the two scanners. Given previous reports of potential sex differences in both the neural basis of dyslexia and the glutamate system, participants' sex was included as a variable of interest in the first two experiments (Giacometti & Barker, 2020).

2.2.Methods

2.2.1. Participants

Participants invited to take part in the functional magnetic resonance spectroscopy study were selected from a larger project focused on the neural noise hypothesis in individuals with dyslexia (OPUS grant, National Science Centre, awarded to professor Katarzyna Jednoróg (2019/35/B/HS6/01763)). Participants were assigned to the dyslexic group based on a clinical diagnosis of dyslexia established during childhood. All participants were right-handed, born at term, had normal or corrected-to-normal vision, no hearing impairments, no history of neurological or psychiatric disorders, and an IQ higher than 80. IQ was assessed using the Polish version of the Abbreviated Battery of the Stanford-Binet Intelligence Scale-Fifth Edition (SB5; Roid et al., 2017).

The study was approved by the ethics committee at the University of Warsaw, Poland (reference number 2N/02/2021). Written informed consent was obtained from all participants or their legal guardians, and participants received financial compensation for their time: 200 PLN for behavioral and fMRI session, 600 PLN for fMRS on the 7T scanner, and 200 PLN for fMRS on the 3T scanner.

Fifty-nine volunteers participated in the fMRS experiment conducted on the 7T scanner. This group included 30 typical readers (14 females, 16 males) aged between 15.38 and 25.71 years (M = 20.89, SD = 3.16) and 29 individuals diagnosed with dyslexia (13 females, 16 males) aged between 15.22 and 25.66 years (M = 20.27, SD = 3.72). Forty volunteers participated in the fMRS experiment conducted on the 3T scanner. This group included 19 typical readers (11 females, 8 males) aged between 15.64 and 25.19 years (M = 21.87, SD = 2.93) and 21 individuals diagnosed with dyslexia (9 females, 12 males) aged between 15.40 and 25.03 years (M = 21.50, SD = 3.40). A total of 37 participants were scanned using both the 3T and 7T scanners. Of these, 22 participants underwent the 3T scanner session first. Among the participants, 18 individuals (11 female and 7 male) were from the control group, while 19 participants (8 female and 11 male) were diagnosed with dyslexia. In the control group, participants scanned on the 7T scanner were aged between 15.38 and 25.71 years (M = 21.89, SD = 3.16), and on the 3T scanner between 15.64 and 25.19 years (M = 21.79, SD = 2.99). In the dyslexic group, participants scanned

on the 7T scanner were aged between 15.22 and 25.66 years (M = 21.52, SD = 3.52), and on the 3T scanner between 15.40 and 25.03 years (M = 21.19, SD = 3.42).

2.2.2. Procedure

All participants took part in an initial behavioral and fMRI session and in at least one fMRS session. During the behavioral session participants completed a battery of paper-pencil reading and reading-related tests to assess their reading skills. This phase lasted approximately one hour. Subsequently, participants underwent the fMRI experiment to accurately determine the coordinates of the reading-sensitive regions for fMRS voxel placement. They were first instructed on the tasks to be performed in the scanner. The fMRI session included a T1-weighted anatomical scan, followed by three runs of the fMRI experiment. The entire fMRI procedure lasted about 40 minutes, and the total session duration was approximately two hours.

On a separate day the participant took part in the fMRS experiment. The session began with task instructions for the fMRS procedure. Participants first underwent a T1-weighted anatomical scan, which was then co-registered with the anatomical scan from the fMRI session. To ensure optimal field homogeneity within the selected voxel, an initial shimming measurement was performed, followed by shimming calibration. The fMRS experiment was initiated when the voxel linewidth was below 20 Hz. The total duration of the fMRS session was approximately two hours, depending on the number of fMRS runs successfully completed (based on shimming results). Brain regions were scanned in a pseudo-random order. On the 7T scanner, an additional anatomical scan was acquired for the brain volumetry.

2.2.3. Reading tests

Participants' reading skills were assessed using a variety of paper-and-pencil tasks (see Glica et al., 2024 for details). Reading skills were measured by counting the number of words and pseudowords participants could read correctly within one minute (Szczerbiński & Pelc-Pękała, 2013). Reading comprehension was assessed with 26 short sentences (e.g., "Lemons are yellow," "A year has seven months") that participants read silently. They indicated whether the sentence was true or false, and the time to complete the task served as the outcome measure.

Additionally, lexical access was determined using Rapid Automatized Naming (RAN) including subtests for naming objects, colors, digits, and letters (Fecenec et al., 2013). Phonological awareness was evaluated using two tasks: a phoneme deletion task (Szczerbiński & Pelc-Pękała, 2013), where participants had to repeat a word after removing a specific phoneme (e.g., "farm" without "f"), and a spoonerisms task (Bogdanowicz et al., 2016), where participants had to switch phonemes or syllables between two words. Orthographic awareness was measured using 28 pairs of pseudowords, where one was written according to Polish spelling rules, and the other was not. Participants identified the correctly written pseudoword. Accuracy divided by time served as the outcome measure (Awramiuk & Krasowicz-Kupis, 2014).

2.2.4. fMRI Experiment

fMRI Task

The aim of the fMRI experiment was to localize the language-sensitive left superior temporal sulcus (STS) and the visual word form area (VWFA) in individual participants.

The fMRI task consisted of three runs, each lasting 5 minutes and 9 seconds. Two runs involved the presentation of visual stimuli, and one involved auditory stimuli. Visual stimuli included three categories: words, consonant strings, and false font strings (BACS font; Vidal et al., 2017). All visual stimuli consisted of 3-4 letter/symbol strings. Auditory stimuli included three categories: words, consonant strings, and backward speech. They were generated using a speech synthesizer (text-to-speech generator) to minimize variability in emphasis.

Stimuli were presented for 800 ms with a 400 ms break in between, forming blocks of 14 stimuli (each block lasting 16.8 seconds). Each run contained 12 stimulation blocks (4 for each category), with a fixation cross displayed for 8 seconds between blocks. Participants performed a 1-back task, where they needed to press a button whenever they saw a stimulus that was the same as the one they had just seen, in order to maintain focus throughout all runs. Each block included 2 to 3 repeated stimuli serving as targets.

The stimuli were presented using Presentation software (Version 20.1, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) running on a PC. Visual stimuli were displayed on a screen positioned behind the MR scanner and viewed via a mirror mounted

above the participants' eyes. Auditory stimuli were delivered through MR-compatible headphones.

fMRI Data Acquisition

fMRI data were acquired using a Siemens 3T Trio system with a 32-channel radiofrequency head coil (two participants were scanned using a 12-channel coil due to technical issues) at Laboratory of Brain Imagining, Nencki Institute of Experimental Biology.

Anatomical data was acquired using a whole-brain 3D T1-weighted image (MPRAGE sequence): inversion time (TI): 1100 ms; GRAPPA parallel imaging with an acceleration factor: PE = 2; voxel resolution: 1 × 1 × 1 mm^3; dimensions: 256 × 256 × 176; acquisition time (TA): 6 minutes 3 seconds.

Functional data was acquired with a whole-brain echo planar imaging (EPI) sequence: echo time (TE): 30 ms; repetition time (TR): 1410 ms; flip angle (FA): 90°; field of view (FOV): 212 mm; matrix size: 92 × 92; multiband acceleration factor: 3; slice thickness: 2.3 mm (60 axial slices); in-plane resolution: 2.3 × 2.3 mm^2.

fMRI data analysis

The data were processed using Statistical Parametric Mapping (SPM12) (Wellcome Trust Centre for Neuroimaging, London, UK) implemented in MATLAB R2020b (The MathWorks Inc., Natick, MA, USA). The preprocessing steps included realignment, coregistration and smoothing. First, all functional images were realigned to the participant's mean image. Then, T1-weighted anatomical images were coregistered to the functional images for each subject. Finally, the fMRI data were smoothed using a 6 mm isotropic Gaussian kernel.

The left superior temporal sulcus (STS) was localized in the native space for each participant as a cluster located in the middle posterior part of the left superior temporal sulcus. This cluster showed significantly higher activation for visual words compared to false font strings and for auditory words compared to backward words (logical AND conjunction), using a threshold of $p < 0.01$ uncorrected. For six participants (DYS: n = 2, CON: n = 4), the threshold was adjusted to $p < 0.05$ uncorrected, and for another six participants (DYS: n = 3, CON: n = 3), the auditory contrast was modified to auditory

words versus fixation cross due to the absence of significant activation for the original contrasts.

The visual word form area (VWFA) was localized in the native space for each participant as a cluster within the occipitotemporal sulcus. This cluster exhibited greater activation for words compared to false font strings, as well as for words compared to fixation cross (logical AND conjunction), using a threshold of $p < 0.01$ uncorrected.

2.2.5. fMRS Experiment

Regions of Interest

The goal of the fMRS study was to measure glutamate concentration changes in brain regions belonging to the reading network. Two key regions associated with reading, the left superior temporal sulcus (STS) and the visual word form area (VWFA), were localized individually for each participant. Additionally, a voxel in a control region, the medial prefrontal cortex (mPFC), not directly involved in reading, was placed based on anatomical localization (Figure 5).

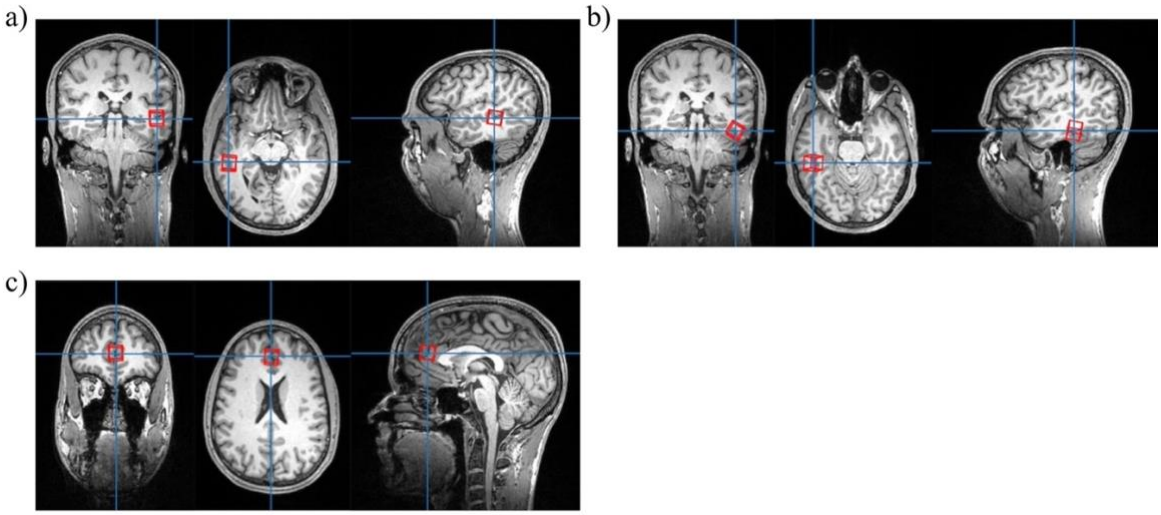

Figure 5. Example of voxel positioning for: a) left superior temporal sulcus (STS), b) visual word form area (VWFA), c) medial prefrontal cortex (mPFC, control region).

fMRS task

During the fMRS task, in contrast to the fMRI localizer, only visual stimuli were used. Each trial consisted of either three Polish words or three false font strings (written in BACS font) presented consecutively. Trials lasted 850 ms, with each word or false font string (consisting of 3–4 letters or symbols) presented for 250 ms, separated by a 50 ms gap (see Figure 6). Each stimulation block consisted of 12 trials (16 MRS averages were acquired), presented in a pseudo-random order (6 trials with words and 6 with false font strings). Each run consisted of 13 stimulation blocks (each lasting 64 seconds) and 13 rest periods (lasting 32–36 seconds, except for the final rest period, which lasted 60 seconds). Additionally, the stimuli were specifically ordered to prevent the repetition of signals of the same type (visual stimulation, delay) within the same cycle of 16-step phase cycling.

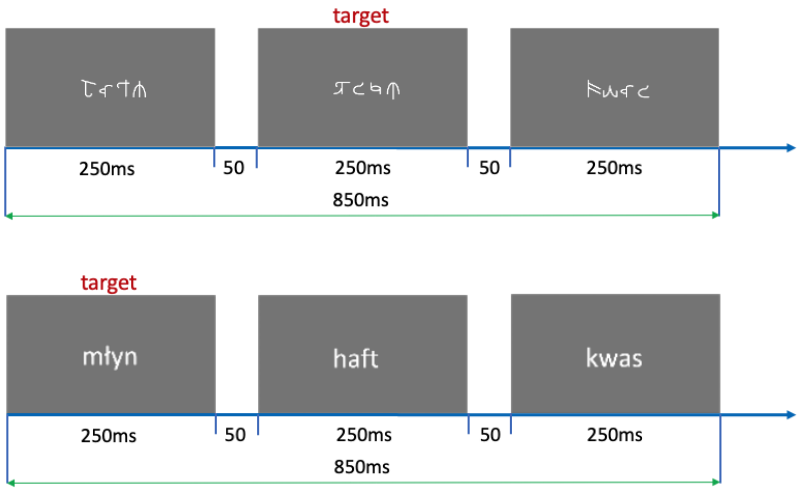

Figure 6. Examples of visual stimuli: false font string (BACS) and words in Polish. Visual stimuli containing letters or symbols with descending features were designated as targets.

To ensure that participants stayed focused, they were instructed to press a button on a response pad whenever they saw a word or false font string containing a descending feature. In each run, 468 stimuli were presented in 156 trials (78 trials with words and 78 trials with false font strings) in a pseudo-random order, with a total of 104 target stimuli. The total duration of each run was 21 minutes and 20 seconds per region of interest.

Similar to the fMRI task, the stimuli were presented using Presentation software (Version 20.1, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) running on a PC. Visual stimuli were displayed on a screen positioned behind the MR scanner and viewed via a mirror mounted above the participants' eyes.

Glutamate dynamics

The fMRS task was designed to measure changes in glutamate concentration over time (see Figure 7). Since the glutamate response function is still a matter of debate, MRS signals were acquired at four time points relative to trial onset: 500 ms, 1000 ms, 3000 ms, and 4500 ms. These delays were chosen because the visual stimuli in this project were complex and were aligned with the repetition time (TR = 4 s) of the experiment.

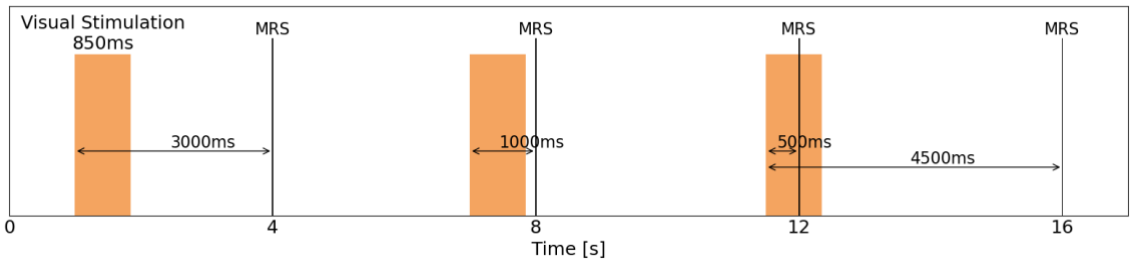

Figure 7. Design of delays between visual stimulation and MRS signal acquisition. MRS signals were acquired with a repetition time (TR) of 4 seconds.

During each stimulation block, 16 MRS averages were collected, two of each type (words with 500 ms delay, words with 1000 ms delay, words with 3000 ms delay, words with 4500 ms delay, false font strings with 500 ms delay, false font strings with 1000 ms delay, false font strings with 3000 ms delay, false font strings with 4500 ms delay). MRS acquisitions at 500 ms and 4500 ms were always collected consecutively, with a repetition time (TR) of 4 seconds.

fMRS Data Acquisition

Scanners

Data were acquired using two scanners: a 7T GE DISCOVERY 950 system with a 32-channel radiofrequency head coil located at the Ultra-High Field Magnetic Resonance Lab, Ecotech-Complex, Maria Curie-Skłodowska University in Lublin and a 3T Siemens Trio system with a 32-channel radiofrequency head coil, previously used for the fMRI study. Data from the 7T scanner were saved in pfile format, while data from the 3T scanner were saved in twix format. Both file types preserved all dimensions (coil channels, individual MRS signals) uncombined, which is critical for functional spectroscopy.

Anatomical Scans

Structural data on the 7T scanner were acquired as a whole-brain 3D T1-weighted image with the following parameters: sequence: 3D-SPGR BRAVO; inversion time (TI): 450 ms; echo time (TE): 2.6 ms; repetition time (TR): 6.6 ms; flip angle: 12°; bandwidth: ±32.5 kHz; ARC acceleration factor: PE = 2; voxel resolution: 1 × 1 × 1 mm^3; dimensions: 256 × 256 × 180.

A T1-weighted anatomical scan on the 3T scanner was acquired using the same parameters as in the fMRI study.

Coregistration of Structural Images

Structural images from the fMRI and fMRS sessions were co-registered using the Statistical Parametric Mapping toolbox (SPM12) (Welcome Trust Center for Neuroimaging, London, UK). This coregistration allowed for accurate positioning of the fMRS voxel based on coordinates from the fMRI localizer for each individual. Voxels were adjusted to ensure they included only brain tissue.

Shimming

Ensuring magnetic field homogeneity within the voxel was a critical step before starting the fMRS experiment. The linewidth of the spectra needed to be less than 20 Hz, which was achieved through a shimming procedure (Landheer & Juchem, 2021). On the 7T scanner, the Famasito 1st and 2nd order shimming algorithm was employed whenever possible; however, the procedure often led to unrealistic settings. Zero- and first-order shims were used when Famasito failed to converge. On the 3T scanner, the 1st and 2nd order FASTMAP algorithm was used to optimize the magnetic field homogeneity.

fMRS Data Acquisition

Functional magnetic resonance spectroscopy was acquired using single-voxel spectroscopy method with a semi-LASER sequence (Deelchand et al., 2021); voxel size: 15 × 15 × 15 mm^3; echo time (TE): 28 ms; repetition time (TR): 4000 ms; data points: 4096; averages: 320; water suppression: VAPOR; reference scans (averages with unsuppressed water without task performance): eight for 7 Tesla, sixteen for 3 Tesla scanners; phase cycling: 16 steps.

2.2.6. fMRS Data Analysis

Basis Set

A basis set is a collection of metabolite response functions required to decompose functional magnetic resonance spectroscopy (fMRS) data into individual components, representing specific metabolite signals. The response functions depend on scanning parameters such as field strength and acquisition sequence. This is a crucial step for quantifying metabolite concentrations from the acquired spectra.

The higher frequency resolution at 7T allows for the detection of a larger number of metabolites. However, it remains uncertain whether including metabolites with theoretically low concentrations improves data analysis. To address this, different metabolite sets were chosen for the 3T and 7T scanners to reduce the risk of misidentifying low-concentration metabolites as glutamate.

For both 3T and 7T scanners, basis-sets were selected to match the main frequency, sequence and echo time used on the scanners.

For the 7T scanner, the basis set included 27 metabolite spectra simulated using FID-A (Simpson et al., 2017) with a script customized for our experiment (Figure 8). The following metabolites were included: Ala, Asc, Asp, Cit, Cr, EtOH, GABA, GPC, GSH, Glc, Gln, Glu, Gly, Ins, Lac, NAA, NAAG, PCh, PCr, PE, Phenyl, Scyllo, Ser, Tau, Tyros, bHB, and bHG. Additionally, synthetic model-based macromolecule spectra provided by FSL-MRS were added.

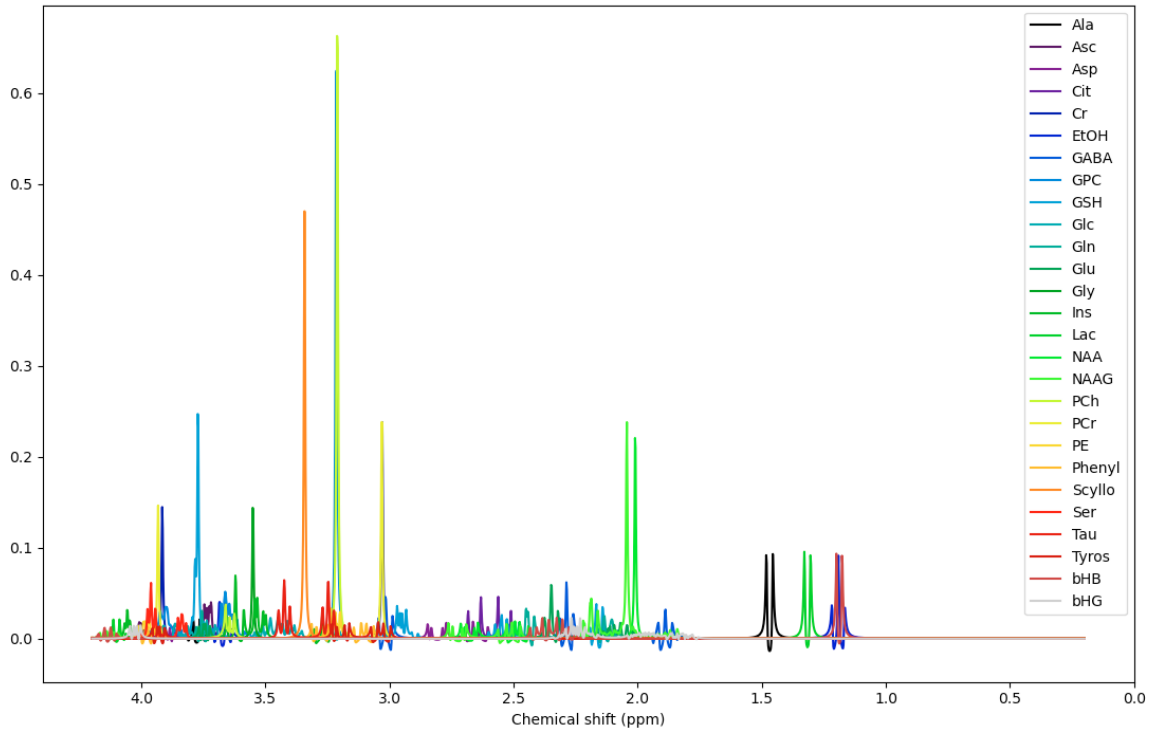

Figure 8. Visualization of the basis set for the 7T scanner without macromolecules added. Visualization generated using the *mrs_tools vis* function.

The basis set for the 3T scanner included 18 metabolite spectra, supplemented with macromolecule spectra using FSL-MRS. The included metabolites were: Ala, Asc, Asp, Cr, GABA, GSH, Glc, Gln, Glu, Ins, Lac, NAA, NAAG, Pcho, PCr, PE, Tau, and sins (see Figure 9). This basis set was not designed specifically for this study and it was publicly available on FSL-MRS page (https://github.com/wtclarke/fsl_mrs/tree/master/fsl_mrs/mmbasis/oldBasisSets). The single broad macromolecule component (Mac) originally present in the basis set was replaced by model-based macromolecular components added using FSL-MRS, in order to match the macromolecular modeling approach used for the 7T data and to facilitate comparison between the two scanners.

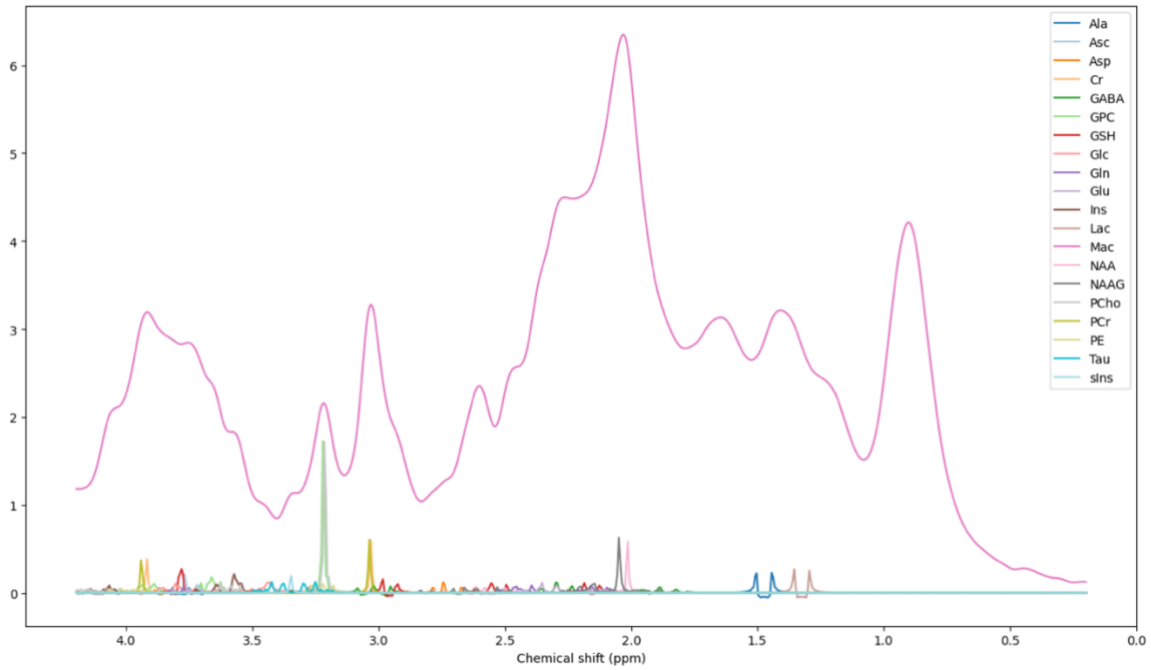

Figure 9. Visualization of the basis set for the 3T scanner without macromolecules added, before removing Mac. Visualization generated using the *mrs_tools vis* function.

Quality Check

Initially, fMRS data were analyzed as conventional magnetic resonance spectroscopy data. All 320 signals acquired during a run were averaged to create a single spectrum. This approach was employed to assess the overall quality of the fMRS data.

Data were analyzed using FSL-MRS (version 2.0.7; Clarke et al., 2021). Raw data stored in pfile and twix formats were converted to NifTI-MRS format using the spec2nii conversion tool available in the FSL-MRS package (https://onlinelibrary.wiley.com/doi/10.1002/mrm.29418, GitHub link).

Basic preprocessing steps were performed automatically using the fsl_mrs_preproc function, including: coil combination, frequency and phase alignment, removal of bad averages, averaging all 320 signals into a single spectrum, eddy current correction, frequency shifting to the reference peak, phase correction.

Subsequently, the preprocessed data were quantified using the appropriate basis set, with macromolecules treated as a separate group. The analysis employed unsuppressed water signals for eddy current correction. Averaged signals from creatine and phosphocreatine

served as internal references for metabolite concentration. Quantification was performed within the 0.2–4.2 ppm range (see Figure 10).

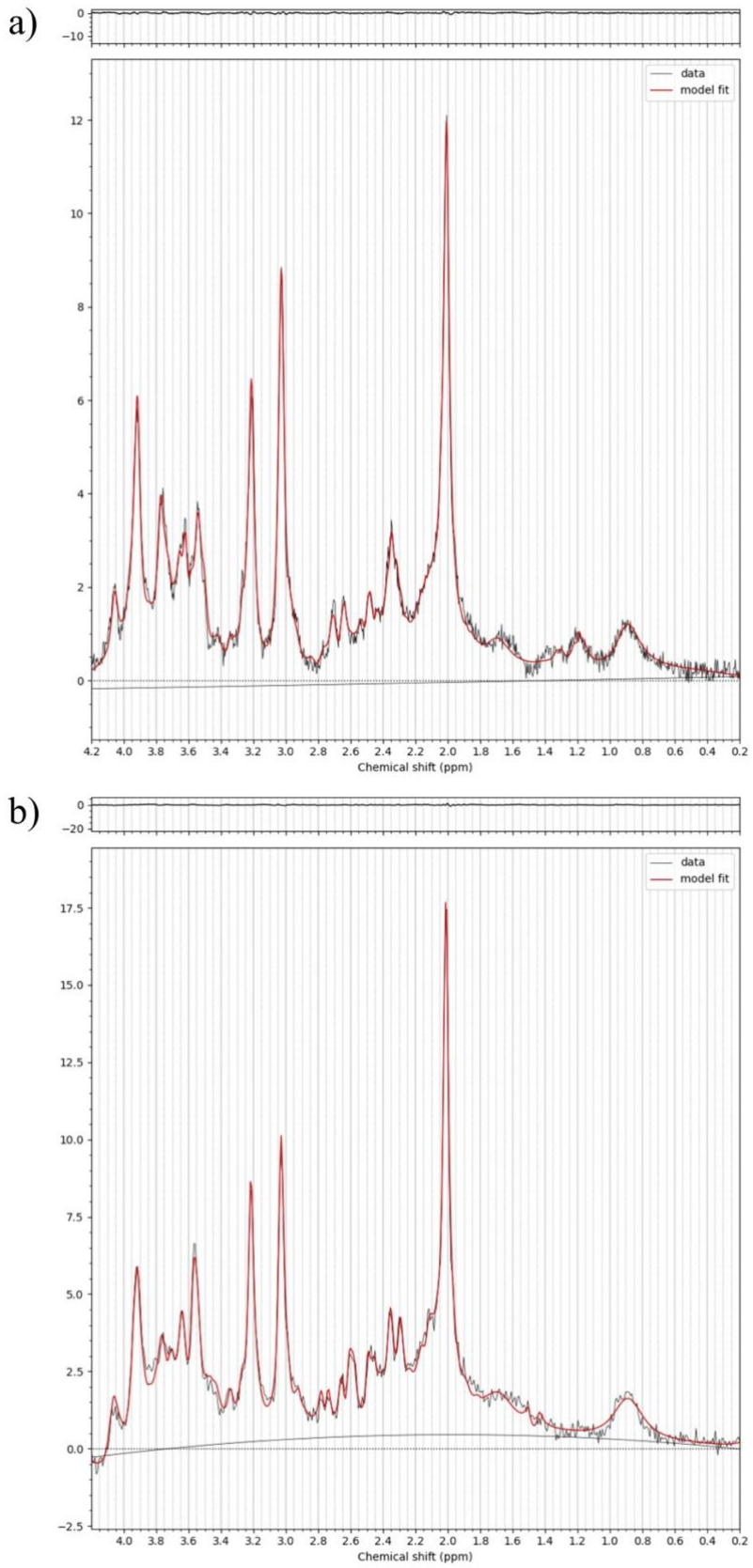

Figure 10. Example of a high-quality single-subject spectrum acquired from the STS a) on a 7T scanner b) on a 3T scanner. The black spectrum represents the original data averaged from 320 signals. The red line

shows the fitted spectrum based on the basis set. Residuals (differences between the original data and the fitted spectrum) are displayed at the top of the figure.

The quality assessment included both quantitative analysis and visual inspection. Spectra with a linewidth higher than thresholds or with poor model fit were excluded from further analysis. The linewidth of metabolite peaks was assessed for both scanners, with different thresholds for each: 20 Hz for 7T and 8.57 Hz for 3T scanner respectively. The difference in linewidth thresholds arises because linewidth expressed in Hz is strongly linked to the Larmor frequency (f_{Larmor}) and the full width at half maximum (FWHM), which can be expressed in both Hz and ppm. To maintain a consistent FWHM in ppm across scanners, the52hresholdds in Hz must differ.

The relationship between FWHM in ppm and Hz is described by the following formula:

$$f_{Larmor} = \gamma B_0$$

$$FWHM[ppm] = \frac{FWHM[Hz]}{f_{Larmor}}$$

where:
γ – gyromagnetic constant for protons ($42.58\,MHz/T$),
B_0 – magnetic field induction

For the 3T scanner ($f_{Larmor} = 127.7\ MHz$) and the 7T scanner ($f_{Larmor} = 298.0\ MHz$), the thresholds were calculated as follows:

$$FWHM_{7T}[Hz] = 20\ Hz$$

$$\frac{FWHM_{7T}[Hz]}{f_{Larmor,7T}} = \frac{FWHM_{3T}[Hz]}{f_{Larmor,3T}}$$

$$FWHM_{3T}[Hz] = \frac{FWHM_{7T}[Hz] \cdot f_{Larmor,3T}}{f_{Larmor,7T}}$$

$$FWHM_{3T}[Hz] = \frac{20\ Hz \cdot 127.7\ MHz}{298.0\ MHz} = 8.57\ Hz$$

This calculation ensures that the FWHM in ppm remains consistent across scanners.

After quality assessment, some spectra were excluded from the analysis (see Figure 11).

On the 7T scanner, 55 out of 57 medial prefrontal cortex scans were included in the analysis. Among them, 29 were from the control group (14 females, 15 males) and 26 from the dyslexic group (13 females, 13 males). In the left superior temporal sulcus region, 49 out of 54 spectra were usable. This group included 28 control participants (12 females, 16 males) and 21 dyslexic participants (12 females, 9 males). For the visual word form area, only 12 out of 42 data sets were suitable for analysis. This included 7 control participants (3 females, 4 males) and 5 dyslexic participants (3 females, 2 males).

On the 3T scanner, 38 out of 40 medial prefrontal cortex scans achieved good quality data. Of these, 18 were from the control group (11 females, 7 males) and 20 from the dyslexic group (9 females, 11 males). In the left superior temporal sulcus, 35 out of 39 spectra were included in the analysis. This group consisted of 17 control participants (10 females, 7 males) and 18 dyslexic participants (11 females, 7 males). For the visual word form area, 9 out of 21 spectra were acceptable. Of these, 5 were from the control group (2 females, 3 males) and 4 from the dyslexic group (1 female, 3 males).

For 34 participants scanned using both 7T and 3T scanners, medial prefrontal cortex scans provided data suitable for analysis. This included 18 control participants (11 females, 7 males) and 16 dyslexic participants (8 females, 8 males). 30 spectra from the left superior temporal sulcus were good quality, with 17 from the control group (10 females, 7 males) and 13 from the dyslexic group (6 females, 7 males). Due to the low quality of the visual word form area data (with only a few datasets meeting the quality criteria), this region was omitted from further analysis.

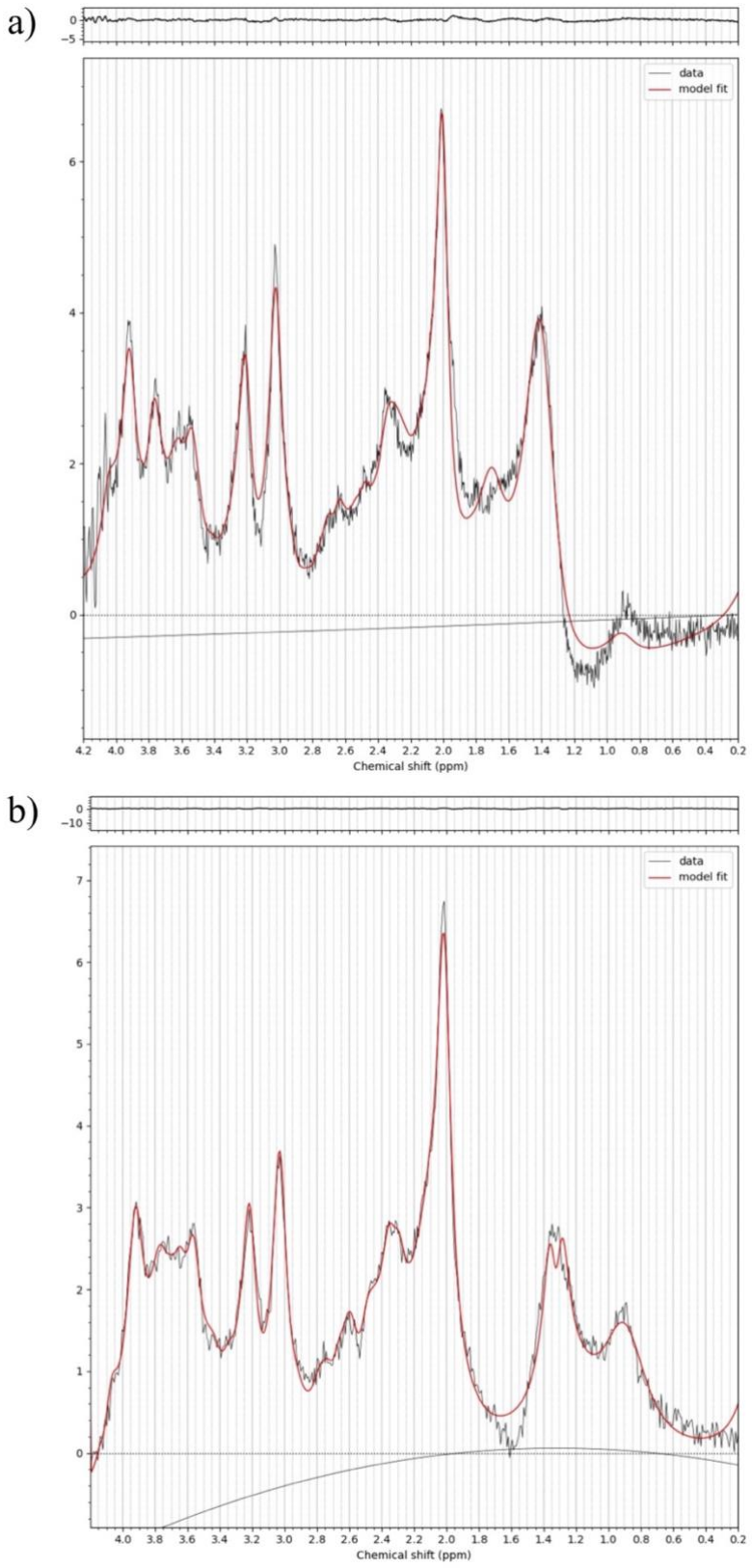

Figure 11. Example of a low-quality single-subject spectrum acquired from the STS, omitted from the analysis, a) on a 7T scanner b) on a 3T scanner. The black spectrum represents the original data averaged from 320 signals. The red line shows the fitted spectrum based on the basis set. Residuals (differences between the original data and the fitted spectrum) are displayed at the top of the figure.

fMRS Preprocessing

After the quality assessment, the data were analyzed using functional magnetic resonance spectroscopy. All preprocessing steps were performed using FSL-MRS (version 2.0.7; Clarke et al., 2021). The data had been previously converted to NifTI-MRS format during the MRS analysis. Data from each region of interest (ROI) was preprocessed separately using the fsl_mrs_preproc function. This function performed the following preprocessing steps: coil combination, frequency and phase alignment, phase correction based on the total creatine (creatine and phosphocreatine) peak, eddy current correction.

All signals (320 signals in total) were preprocessed together before being divided into data packages to avoid potential glutamate changes that could arise from differences in the preprocessing steps, rather than from the actual glutamate response to the visual stimulation.

Importantly, no MRS scans were averaged at this stage (the *fmrs=True* option was used in the fsl_mrs_preproc function). Additionally, the "unlike" option, which automatically removes bad averages, was deactivated. This choice allowed for greater control over which signals were excluded from the dataset.

Voxel Segmentation

To determine the voxel composition – specifically, the percentage of white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), the svs_segmentation tool was used, with the results of fsl_anat as input. Voxel segmentation was performed on structural images acquired from the 3T scanner for both datasets (3T and 7T). Anatomical images from the 3T scans were coregistered to the corresponding 7T images using SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK). This approach was chosen because T1-weighted images from the 3T scanner had fewer artifacts and less intensity bias in the temporal lobe compared to 7T anatomical images. These differences are due to the higher magnetic field inhomogeneity associated with the 7T scanner.

fMRS Data Analysis – Conventional Averaged Approach

The data were analyzed by dividing signals into groups based on the type of stimulation. This approach improved the signal-to-noise ratio (SNR). The number of MRS signals in each group was equalized to minimize errors caused by differences in data quality. More averaged MRS signals led to better spectra quality, which could influence fitting and, consequently, metabolite concentration results.

Grouping by Stimulation Type

Initially, MRS signals were divided into three groups: words in Polish (78 MRS acquisitions), false font strings (78 MRS acquisitions), rest (78 MRS acquisitions). For both visual stimulations, MRS signals acquired with time delays of 500 ms, 1000 ms, and 3000 ms were analyzed together as one group (see Figure 12).

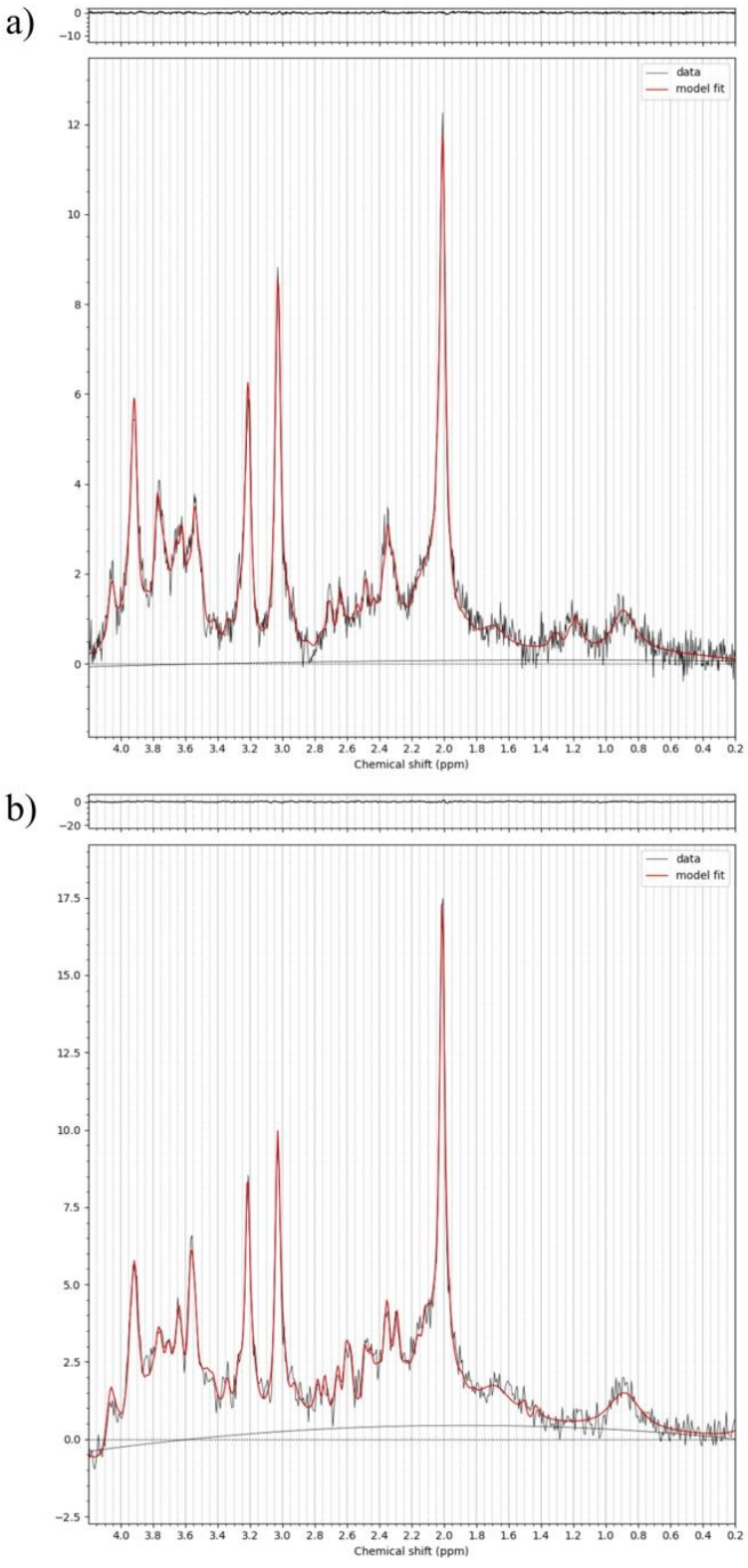

Figure 12. Example of a single-subject spectrum acquired from the STS a) on a 7T scanner b) on a 3T scanner. The black spectrum represents the original data averaged from 78 signals acquired after words stimulation. The red line shows the fitted spectrum based on the basis set. Residuals (differences between the original data and the fitted spectrum) are displayed at the top of the figure.

Grouping by Delay

To investigate differences between delays, the data were further divided into 8 groups**:** words with 500 ms delays (26 MRS acquisitions), words with 1000 ms delays (26 MRS acquisitions), words with 3000 ms delays (26 MRS acquisitions), words with 4500 ms delays (26 MRS acquisitions), false font strings with 500 ms delays (26 MRS acquisitions), false font strings with 1000 ms delays (26 MRS acquisitions), false font strings with 3000 ms delays (26 MRS acquisitions), false font strings with 4500 ms delays (26 MRS acquisitions) (see Figure 13).

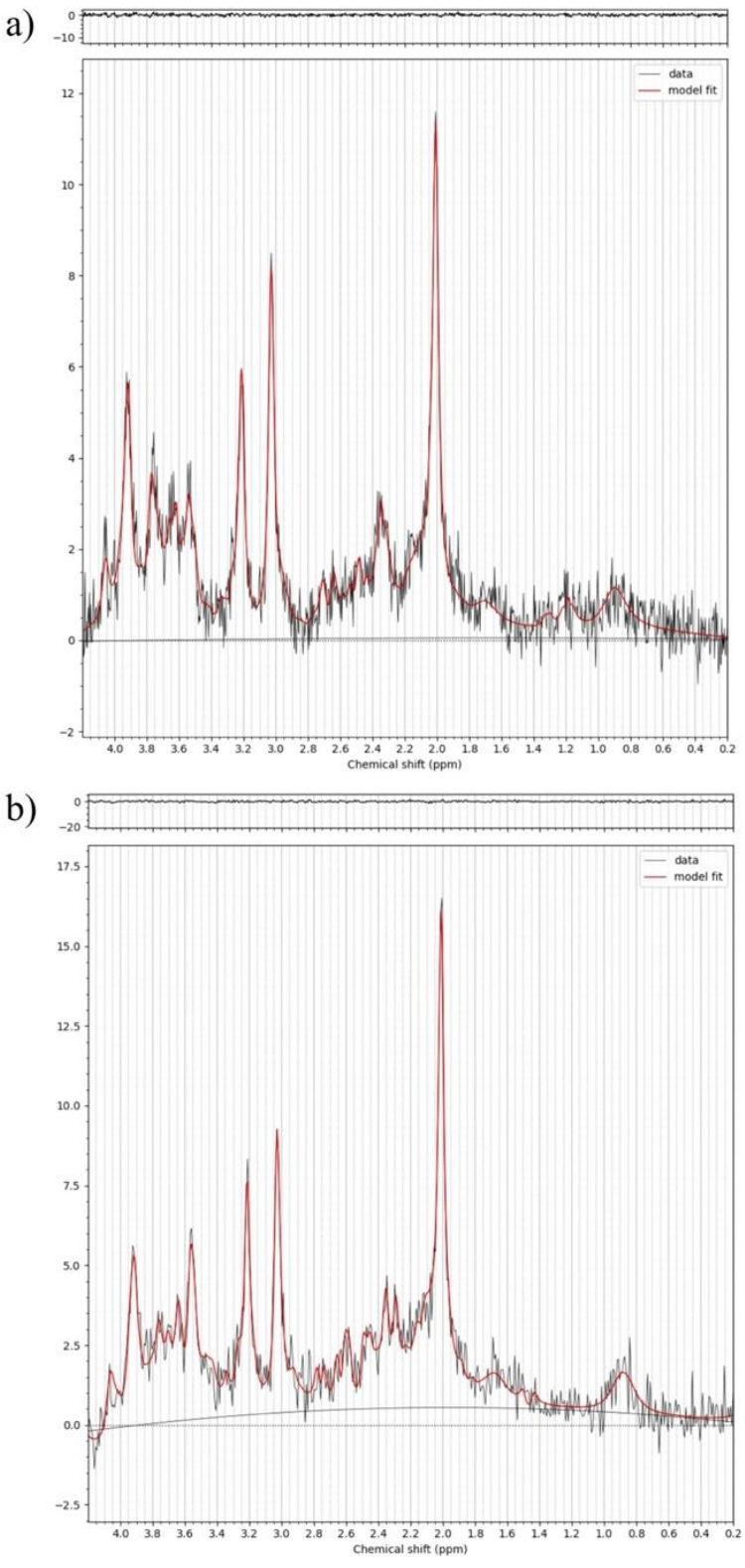

Figure 13. Example of a single-subject spectrum acquired from the STS a) on a 7T scanner b) on a 3T scanner. The black spectrum represents the original data averaged from 26 signals acquired after words stimulation with 500 ms delay. The red line shows the fitted spectrum based on the basis set. Residuals (differences between the original data and the fitted spectrum) are displayed at the top of the figure.

Detection of Unlike Signals

After fMRS preprocessing, unlike signals were identified using the *identifyUnlikeFIDs* function (GitHub repository) with a standard deviation limit of 1.96 (default setting). The output was a list of indexes corresponding to signals potentially affected by movement or other artifacts.

Before data averaging, signals identified as unlike were excluded. Subsequently, the number of MRS signals in each group was evaluated. Data sets in which more than 20% of signals were excluded in any group were generally removed from further analysis. However, one data set acquired from the STS region on the 7T scanner had 6 unlike signals out of 26 in a single group (23%). While this exceeds the standard exclusion threshold, the participant was included in the analysis because the remaining groups within the data set met the quality criteria, and excluding this participant would result in an unnecessary data loss.

Ultimately, no data set was excluded from the analysis due to an excessive number of signals identified as unlike.

Fitting

MRS signals from functional magnetic resonance spectroscopy divided into groups were fitted using FSL-MRS (version 2.1.20) with default settings. MRS signals accumulated in groups were averaged after excluding unlike signals. Metabolites were treated as separate components, with each fitted individually using its corresponding basis set, while all macromolecules were treated together as a single group. The unsuppressed water signals, averaged across all acquisitions, were used as a reference for quantification. The same reference file was applied to all groups.

fMRS Data Analysis – Mixed Dynamic-Averaged Approach

Especially for this project, an innovative approach was proposed that combines the advantages of the traditional fMRS data analysis method, which relies on averaged MRS signals acquired with the same type of stimulation, with the dynamic approach increasingly used in fMRS data analysis. This is the first time this idea has been applied.

In the dynamic model, multiple spectra can be provided and linked using an arbitrary model. A reduced set of model parameters, such as the concentration of one or more metabolites, linewidth, etc., can vary, while other parameters remain fixed within the common model.

The dynamic method is based on a general linear model (GLM), where a design matrix is used to predict and describe factors influencing individual fMRS signals, such as specific stimulations, BOLD effects, and other predictable variables. The advantage of combining averaged groups with GLM fitting lies in two main benefits: the significantly higher signal-to-noise ratio (SNR) of the averaged groups and the substantial simplification of the design matrix, which would otherwise be highly complex for this project's paradigm.

While the dynamic approach allows for the simultaneous analysis of all groups within a single fitting procedure, it also enables researchers to define which metabolite concentrations are expected to differ in response to stimuli (calculated separately for each group) and which remain constant across the entire experiment. This flexibility reduces the risk of misinterpreting spectral overlaps from other metabolites as actual concentration changes during the experiment. By fixing the concentrations of metabolites whose spectral overlap may interfere with the target metabolite (e.g., glutamate) and whose changes in response to stimulation are not expected, the dynamic approach minimizes fitting errors and provides a clearer picture of the metabolite of interest.

Additionally, performing a single fitting procedure for all groups prevents potential inconsistencies in metabolite concentrations that could arise from separate quantification in FSL-MRS, where initial conditions for the analysis may differ between groups. These variations in starting parameters, such as linewidth, frequency shifts, or baseline characteristics, can lead to differences in metabolite quantification and make comparisons between groups less reliable. By fitting all groups simultaneously, the dynamic approach ensures that the same reference conditions are used, reducing systematic biases and improving the robustness of the analysis.

Dynamic Fitting of averaged groups

At the beginning of the analysis, data averaged into groups (based on the type of visual stimulation and the time delays between stimulation and MRS acquisition, already prepared for averaged approach analysis) were merged into a NifTI-MRS file containing eight averaged signals. The signals were ordered as follows: averaged group of words with a 500 ms delay, averaged group of words with a 1000 ms delay, averaged group of words with a 3000 ms delay, averaged group of words with a 4500 ms delay, averaged group of false font strings (BACS) with a 500 ms delay, averaged group of false font strings (BACS) with a 1000 ms delay, averaged group of false font strings (BACS) with a 3000 ms delay, averaged group of false font strings (BACS) with a 4500 ms delay. This merged file is treated as spectrum with eight MRS signals. Dynamic fitting was performed using FSL-MRS (version 2.1.20), with macromolecules treated as a separate group. Other parameters were set to default.

Design matrix

Despite the fact that none of the metabolites were fitted dynamically, the software required the use of the design matrix to run the fitting. The design matrix loaded contained nine columns. The first eight columns corresponded to the eight groups, with each column containing ones for the respective group and zeros for others. The ninth column was the intercept, consisting entirely of ones.

Configuration file

The configuration file is a Python module (py file) where researchers define expectations how metabolites and other parameters vary during the run. This file specifies whether a parameter: remains constant throughout the experiment (Fixed), changes dynamically across groups based on a General Linear Model (GLM), where values are calculated for each column in the design matrix to determine how much they contribute to the signal (Dynamic) or varies for each individual signal (Variable). The configuration file also allows for setting constraints on parameter ranges (e.g., positive values for concentrations), and implementing custom dynamic models tailored to specific hypotheses.

In this study, glutamate was set as variable to allow its concentration to vary across eight groups, while all other metabolites were fixed at constant values throughout the experiment.

Additionally, boundaries for metabolite concentrations, as well as the gamma (responsible for the Lorentzian part of the spectrum) and sigma (responsible for the Gaussian part of the spectrum) parameters were defined from 0 to infinity to ensure physiologically meaningful results.

Spectral Broadening – BOLD Effect Correction

Changes in the calculated glutamate concentration may be influenced by the blood oxygenation level dependent (BOLD) effect. Oxygenated blood flows into brain regions involved in the task, which can modify the local magnetic field. As a result, the magnetic resonance spectrum may become narrower, potentially leading to incorrect metabolite quantification. To ensure that the observed changes are due to actual variations in glutamate concentration rather than being influenced by the BOLD effect, the data were reanalyzed after equalizing spectral broadening across groups. This procedure was introduced to isolate the contribution of the BOLD effect and confirm that the observed changes in glutamate concentration reflected true metabolic alterations.

The BOLD correction procedure was conducted using a custom Python script designed to utilize the FSL-MRS functions available in open-source software. This approach, based on spectral broadening, has become commonly used, however, researchers often rely on custom scripts and propose different details due to the lack of standardization.

The first step in this procedure was to measure the full width at half maximum (FWHM) of creatine (Cr) using the FSL-MRS function. The group with the largest Cr FWHM was selected as the target, and a spectral broadening procedure was applied. The spectra were incrementally broadened through a stepwise apodization procedure until the desired FWHM was reached. The FWHM of the resulting spectra was compared to the target value, and the spectrum closest to the target was selected. This process was repeated until all groups had similar Cr FWHM values.

The broadening procedure was conducted separately for two different data groupings: by stimulation type, with three groups (Polish words, false font strings, and rest), and by stimulation and delays, resulting in eight groups (words with a 500 ms delay, words with a 1000 ms delay, words with a 3000 ms delay, words with a 4500 ms delay, false font

strings (BACS) with a 500 ms delay, false font strings (BACS) with a 1000 ms delay, false font strings (BACS) with a 3000 ms delay, false font strings (BACS) with a 4500 ms delay).

All data packages used in this step were previously created for the conventional averaged approach. Finally, all analyses were repeated to determine whether the observed changes in glutamate concentration could be contaminated by the BOLD effect rather than reflecting true metabolic changes.

2.3. Statistical analyses

Data from the 3T and 7T scanners were processed as independent datasets in separate analyses. All statistics were performed using JASP (most of them on version 0.19, while some Bayesian ANOVA were conducted on version 0.95 for technical reasons).

2.3.1. Behavioral and reading-related tasks

Behavioral and reading-related tasks were analyzed using separate ANOVAs, with group (dyslexic, control) and sex (female, male) as between-subject factors. All participants scanned using 7T or 3T scanners (separately for each scanner) were included in this analysis, without excluding participants with low-quality spectra.

2.3.2. fMRS analysis

Analyses were conducted separately for the left superior temporal sulcus and the medial prefrontal cortex. The visual word form area was omitted from the analysis due to an insufficient number of high-quality data.

All fMRS analyses were performed using repeated measures ANOVA, with group (dyslexic, control) and sex (female, male) as between-subject factors. Age and gray matter volume (GMV) were included as covariates to control for potential confounding effects.

The data, divided into groups based on stimulation, were analyzed with the type of stimulation (words, false font strings, rest) as the within-subject factor. Additionally, data grouped based on delays and the dynamic-averaged approach were analyzed with the type of visual stimulation (words, false font strings) and delay (500 ms, 1000 ms, 3000 ms, 4500 ms) as within-subject factors.

To follow up on significant interaction effects observed in the ANOVA, planned pairwise contrasts were conducted within the ANOVA framework using custom contrast weights. Specifically, comparisons were performed (a) between the two stimulation types at each level of delay, and (b) between the delays within each stimulation type. Contrasts were specified only for theoretically relevant pairs of conditions and did not include cross-condition comparisons (e.g., words at 500 ms vs false font strings at 1000 ms), as these were not meaningful in the context of the study design. Contrast weights of 1 and -1 were applied to the conditions of interest, and the corresponding t, p, and effect size (d) values were reported. Effect size was interpreted as small ($d \leq 0.2$), medium ($0.2 < d < 0.8$), and large ($d \geq 0.8$).

Bayes Factors from the Bayesian ANOVA for the inclusion of specific effects (BF_{incl}) were computed using the 'across matched model' method. A BF_{incl} value greater than 3 was interpreted as supporting the alternative hypothesis, and a value less than $\frac{1}{3}$ as supporting the null hypothesis. Values between $\frac{1}{3}$ and 3 were considered insufficient evidence to favor either hypothesis (Keysers et al., 2020, Kelter, 2020). Bayes Factors were calculated for main effects and their interactions, but not for covariates or interactions involving covariates.

2.3.3. 7T and 3T scanners comparison

For parameters describing spectra quality, a paired samples t-test was conducted, with t, p, and effect size (d) reported. Cohen's d effect size was interpreted as small ($d \leq 0.2$), medium ($0.2 < d < 0.8$), and large ($d \geq 0.8$). When the assumption of normality was violated (Shapiro-Wilk test: $p < 0.05$), a nonparametric Wilcoxon signed-rank test was used instead of the parametric t-test. For nonparametric tests, the effect size r was calculated and interpreted as small ($|r| \sim 0.1$), medium ($|r| \sim 0.3$), and large ($|r| \geq 0.5$).

3. Results

3.1. Experiment 1 – fMRS on 7T scanner

3.1.1. Behavioral results

Functional magnetic resonance spectroscopy (fMRS) was conducted on a 7 Tesla scanner with 59 participants—29 diagnosed with dyslexia and 30 typical readers. The groups were matched for sex, age, and family socio-economic status, based on parents' educational level.

Volunteers diagnosed with dyslexia had higher scores on the Polish version of the Adult Reading History Questionnaire (ARHQ-PL), indicating a higher risk for dyslexia. Participants in both groups had intelligence quotients (Iqs) within the normal intellectual range, although the dyslexic group scored lower, including the nonverbal subscale (see Table 2).

The control and dyslexic groups also differed in reading ability. Typical readers responded faster in rapid automatized naming and scored higher on phonological awareness tasks. There were no sex differences in the demographic characteristics or behavioral results. Additionally, the interaction between sex and group was not statistically significant for reading-related tasks (see Table 3).

Table 2. Descriptive statistics for demographic characteristics separately for females and males in dyslexic and control groups. For all comparisons, F-statistics, p-values (in brackets), and the partial eta squared (η^2_p) are provided.

	DYS F (n = 13)		CON F (n = 14)		DYS M (n = 16)		CON M (n = 16)		group $F(1,57)$	η^2_p	sex $F(1,57)$	η^2_p	group*sex $F(1, 57)$	η^2_p
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*						
Age	20.978	4.055	21.429	3.530	19.698	3.451	20.424	2.835	0.423 (0.518)	0.008	1.597 (0.212)	0.028	0.023 (0.880)	4.194×10^{-4}
Mother's education (years)	16.500	2.693	16.071	2.814	17.156	4.032	17.250	1.983	0.046 (0.831)	8.315×10^{-4}	1.375 (0.246)	0.024	0.111 (0.740)	0.002
Father's education (years)	15.333[a]	3.869[a]	16.643	4.125	17.000	2.608	16.563	3.032	0.235[b] (0.630)	0.004	0.776[b] (0.382)	0.014	0.941[b] (0.336)	0.017
IQ	102.154	14.577	111.143	7.655	102.500	11.425	114.625	9.952	13.278 (< .001)	0.194	0.437 (0.512)	0.008	0.293 (0.591)	0.005
Nonverbal IQ (scaled score)	10.000	2.887	11.571	2.277	10.375	3.074	12.813	2.040	8.702 (0.005)	0.137	1.414 (0.239)	0.025	0.406 (0.527)	0.007
ARHQ-PL	53.923	9.613	25.214	5.162	48.375	11.899	23.688	7.409	130.307 (< .001)	0.703	2.288 (0.136)	0.040	0.739 (0.394)	0.013

Note. CON – control group; DYS – dyslexic group; F – females; M – males; ARHQ-PL – Polish version of the Adult Reading History Questionnaire.
Boldface indicates statistical significance at $p < .05$ level.
[a] n = 12 (one participant did not provide information about the father's education)
[b] $F(1,56)$

Table 3. Behavioral results from reading and reading-related tasks. For all comparisons, F-statistics, p-values (in brackets), and the partial eta squared (η^2_p) are provided.

	DYS F (n = 13)		CON F (n = 14)		DYS M (n = 16)		CON M (n = 16)		group $F(1,57)$	η^2_p	sex $F(1,57)$	η^2_p	group*sex $F(1,57)$	η^2_p
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*						
words/min	102.923	29.033	137.500	12.451	109.000	16.042	133.500	12.628	38.201 (< .001)*	0.410	0.047 (0.829)	8.577×10^{-4}	1.111 (0.296)	0.020
pseudowords/min	54.769	16.813	85.357	16.570	57.063	13.031	82.688	17.083	45.787 (< .001)*	0.454	0.002 (0.964)	3.732×10^{-5}	0.357 (0.553)	0.006
reading comprehension (sec)	69.462	29.993	40.929	6.810	65.875	20.513	46.500	7.439	24.909 (< .001)*	0.312	0.043 (0.837)	7.768×10^{-4}	0.910 (0.344)	0.016
RAN objects (sec)	32.462	4.521	28.000	5.533	33.625	6.302	28.813	2.949	12.637 (< .001)*	0.187	0.574 (0.452)	0.010	0.018 (0.893)	3.289×10^{-4}
RAN colors (sec)	35.000	4.950	28.929	3.174	36.813	7.287	31.875	3.500	17.355 (< .001)*	0.240	3.243 (0.077)	0.056	0.184 (0.670)	0.003
RAN digits (sec)	21.154	4.652	17.000	3.113	19.000	4.099	16.500	2.449	12.242 (< .001)*	0.182	1.947 (0.168)	0.034	0.756 (0.388)	0.014
RAN letters (sec)	23.385	4.194	18.786	1.929	22.250	3.890	20.000	2.757	15.719 (< .001)*	0.222	0.002 (0.963)	3.867×10^{-5}	1.849 (0.179)	0.033
phoneme deletion (% correct)	68.638	30.401	94.505	5.377	77.644	26.656	93.751	7.414	15.446 (< .001)*	0.219	0.597 (0.443)	0.011	0.835 (0.365)	0.015
spoonerisms phonemes (% correct)	51.098	41.634	88.776	12.441	43.750	35.174	86.159	7.591	30.555 (< .001)*	0.357	0.473 (0.494)	0.009	0.107 (0.745)	0.002

spoonerisms syllables (% correct)	43.590	35.051	74.999	25.943	41.666	31.031	79.166	15.515	22.996 (< .001)*	0.295	0.024 (0.877)	4.426×10^{-4}	0.180 (0.673)	0.003
orthographic awareness (accuracy/time)	0.369	0.155	0.606	0.119	0.327	0.139	0.504	0.131	33.779 (< .001)*	0.380	4.118 (0.047)	0.070	0.702 (0.406)	0.013

Note. CON – control group; DYS – dyslexic group; F – females; M – males; RAN – rapid automatized naming. Boldface indicates statistical significance at $p < .05$ level.
* Significance after Bonferroni correction for 11 planned comparisons for reading-related tasks; threshold for statistical significance after correction was set at p = 0.00455;

3.1.2. Task During fMRS Scanning

Participants were instructed to respond by clicking on the response pad every time they saw a target word on the screen. The accuracy of target recognition was assessed by calculating the percentage of correct responses. Data were analyzed only from good-quality scans. One female participant from the control group was omitted from the analysis due to the lack of saved task logs.

The total number of responses (correct + incorrect) was similar in both brain regions. In the control region, the mean number of responses was 115.296 (SD = 47.651, Min = 45, Max = 415), while in the left superior temporal sulcus, the mean was 112.146 (SD = 48.945, Min = 48, Max = 418).

Brain regions were scanned in pseudorandom order. The medial prefrontal cortex was scanned first 28 times (26 for valid scans; one participant did not have saved logs, and one participant clicked the pad 415 times), second 5 times, and third 22 times. The left superior temporal sulcus was scanned first 21 times, second 26 times (24 for valid scans; one participant did not have saved logs, and one participant clicked the pad over 400 times), and third only 3 times.

The analysis of task accuracy percentage during the scan was conducted using univariate ANOVA. There were no statistically significant effects of group, sex, or interaction between group and sex in either the medial prefrontal cortex or the left superior temporal sulcus (see Table 4).

Table 4. Behavioral results from the task performed during the scan for data that meet the quality criteria. For all comparisons, F-statistics, p-values (in brackets), and the partial eta squared (η^2_p) are provided.

Target recognition (% correct)	DYS F		CON F		DYS M		CON M		group $F(1,53)$ $F(1,47)$	η^2_p	sex $F(1,53)$ $F(1,47)$	η^2_p	group*sex $F(1,53)$ $F(1,47)$	η^2_p
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*						
Medial prefrontal cortex	68.121	7.415	69.601	11.401	62.056	16.373	68.333	11.104	1.412 (0.240)	0.027	1.262 (0.267)	0.025	0.540 (0.466)	0.011
Left superior temporal sulcus	65.385	15.341	65.297	17.659	60.150	12.183	70.853	6.500	1.894 (0.176)	0.041	0.002 (0.967)	3.938×10^{-5}	1.957 (0.169)	0.043

3.1.3. fMRS Results

3.1.3.1. Quality Assessment

The quality of the data, divided into smaller packages based on stimulation type or delay, was assessed using parameters calculated by the FSL-MRS software. To assess quality, the signal-to-noise ratio (SNR), the percentage of Cramer-Rao lower bound (%CRLB), and the full width at half maximum (FWHM) for glutamate and NAA were measured. Analyzed packages were checked against quality criteria, %CRLB for glutamate cannot be higher than 20% to accurately calculate its concentration. No packages had to be excluded. The maximum %CRLB for glutamate in the medial prefrontal cortex was 7.079% for the package with words stimulation at 3000 ms, and the highest %CRLB value in the left superior temporal sulcus was 13.301% for words stimulation at 500 ms. The results are presented in Table 5.

Table 5. Quality parameters of the data, grouped based on stimulation type and delay.

Packages type	Max # signals	Glutamate				NAA			
		Mean	Sd	Min	Max	Mean	Sd	Min	Max
Signal-to-noise ratio (SNR) in mPFC (55 participants)									
Words	78	13.678	2.591	8.421	20.979	35.714	10.144	18.354	64.502
Bacs	78	13.767	2.578	8.942	20.716	35.967	10.153	17.326	65.083
Rest	78	13.657	2.520	9.048	20.764	36.064	10.127	17.337	65.512
words 500 ms	26	8.049	1.451	4.775	11.420	20.955	5.652	10.780	35.521
words 1000 ms	26	8.010	1.423	5.150	11.803	21.004	5.705	11.624	36.303
words 3000 ms	26	7.948	1.545	5.188	12.138	20.768	5.943	9.964	36.979
words 4500 ms	26	7.900	1.353	5.145	10.882	20.937	5.681	10.292	34.921
bacs 500 ms	26	8.029	1.423	5.219	11.823	20.885	5.675	9.984	36.502
bacs 1000 ms	26	8.067	1.528	5.289	11.682	21.116	5.910	9.529	37.260
bacs 3000 ms	26	8.210	1.429	5.744	11.662	21.228	5.830	10.535	36.883
bacs 4500 ms	26	8.023	1.484	5.254	12.186	21.104	5.909	10.516	38.110
Signal-to-noise ratio (SNR) is STS (49 participants)									
Words	78	10.749	3.013	5.438	19.281	38.745	11.293	18.336	75.250
Bacs	78	10.859	3.031	5.186	18.885	38.835	11.658	18.356	74.775
Rest	78	10.748	2.983	5.439	17.960	38.637	11.370	17.174	70.614
words 500 ms	26	6.240	1.780	3.311	10.711	22.475	6.543	11.062	41.651
words 1000 ms	26	6.249	1.737	3.186	10.754	22.562	6.544	10.602	41.833
words 3000 ms	26	6.234	1.650	3.296	10.911	22.555	6.345	10.374	42.523
words 4500 ms	26	6.326	1.784	3.080	11.159	22.558	6.633	10.647	42.442

bacs 500 ms	26	6.310	1.733	3.286	10.643	22.558	6.696	10.107	42.322
bacs 1000 ms	26	6.277	1.732	2.784	10.742	22.536	6.722	10.538	43.356
bacs 3000 ms	26	6.290	1.705	3.237	10.566	22.496	6.605	11.072	42.162
bacs 4500	26	6.274	1.718	3.164	10.065	22.529	6.596	10.607	40.644
%CRLB in mPFC (55 participants)									
Words	78	2.733	0.500	1.894	4.375	2.078	0.498	1.044	3.420
Bacs	78	2.737	0.486	1.943	3.816	2.118	0.549	1.049	3.510
Rest	78	2.798	0.597	1.969	4.820	2.217	0.689	1.128	4.668
words 500 ms	26	4.009	0.930	2.598	6.595	3.085	0.943	1.673	5.581
words 1000 ms	26	3.981	0.869	2.564	6.023	3.115	0.887	1.797	5.325
words 3000 ms	26	4.081	1.052	2.529	7.079	3.277	1.167	1.620	7.028
words 4500 ms	26	4.035	0.880	2.673	6.432	3.073	0.883	1.715	5.387
bacs 500 ms	26	3.969	0.876	2.551	6.477	3.061	0.882	1.775	5.338
bacs 1000 ms	26	3.942	0.842	2.570	6.302	2.969	0.796	1.607	5.055
bacs 3000 ms	26	3.978	0.943	2.625	6.664	3.110	0.992	1.372	5.881
Bacs 4500 ms	26	4.090	0.972	2.533	6.539	3.251	1.019	1.754	6.414
%CRLB in STS (49 participants)									
Words	78	3.776	0.966	2.501	7.379	2.317	0.872	1.115	5.428
Bacs	78	3.761	0.960	2.435	6.711	2.291	0.762	1.095	4.652
Rest	78	3.761	0.954	2.437	6.691	2.308	0.653	1.134	4.481
words 500 ms	26	5.406	1.848	3.267	13.301	3.422	1.569	1.346	9.627
words 1000 ms	26	5.137	1.389	3.223	8.834	3.012	1.134	1.349	6.823
words 3000 ms	26	5.212	1.431	3.248	9.468	3.134	1.225	1.341	7.467
words 4500 ms	26	5.212	1.368	3.220	8.630	3.141	0.931	1.354	5.106
bacs 500 ms	26	5.249	1.460	3.267	9.212	3.200	1.277	1.326	8.341
bacs 1000 ms	26	5.180	1.410	3.227	9.013	3.120	1.108	1.310	7.557
bacs 3000 ms	26	5.245	1.499	3.267	10.046	3.294	1.380	1.366	8.545
bacs 4500 ms	26	5.168	1.476	3.218	9.368	3.151	1.155	1.373	6.881
FWHM in mPFC (55 participants)									
Words	78	26.422	3.437	20.806	34.103	13.807	3.394	8.331	20.573
Bacs	78	26.407	3.444	20.952	34.113	13.872	3.504	8.337	20.778
Rest	78	26.313	3.286	20.857	33.919	13.930	3.610	8.314	20.854
words 500 ms	26	26.311	3.340	20.877	34.278	13.842	3.471	8.313	20.756
words 1000 ms	26	26.266	3.236	20.646	33.145	13.845	3.362	8.339	19.871
words 3000 ms	26	26.185	3.179	20.925	33.749	13.905	3.533	8.366	21.303
words 4500 ms	26	26.408	3.448	20.865	34.439	13.947	3.546	8.354	21.355
bacs 500 ms	26	26.351	3.374	21.014	33.395	13.811	3.450	8.373	20.659
bacs 1000 ms	26	26.439	3.556	20.923	35.083	13.910	3.553	8.425	21.833
bacs 3000 ms	26	26.384	3.390	20.962	34.232	13.822	3.494	8.214	20.757

bacs 4500 ms	26	26.262	3.267	20.889	33.953	13.862	3.481	8.234	20.823
FWHM in STS (49 participants)									
Words	78	26.313	2.779	21.801	31.974	13.460	2.894	8.926	20.868
Bacs	78	26.371	2.841	21.770	32.880	13.456	2.786	8.827	19.806
Rest	78	26.334	2.784	21.875	32.898	13.486	2.792	8.930	19.846
words 500 ms	26	26.266	2.766	21.714	31.982	13.498	2.935	8.967	20.474
words 1000 ms	26	26.374	2.809	21.709	32.820	13.369	2.752	8.827	19.839
words 3000 ms	26	26.291	2.743	21.652	33.099	13.417	2.842	8.967	20.014
words 4500 ms	26	26.332	2.729	21.692	32.151	13.341	2.644	8.912	19.299
bacs 500 ms	26	26.293	2.790	21.843	32.219	13.394	2.735	8.903	19.240
bacs 1000 ms	26	26.381	2.889	21.737	32.802	13.412	2.789	8.789	19.726
bacs 3000 ms	26	26.307	2.740	21.676	32.359	13.519	2.924	8.750	21.020
bacs 4500 ms	26	26.316	2.785	21.629	32.126	13.402	2.744	8.814	19.029

3.1.3.2. Data grouped by stimulation type

Medial prefrontal cortex

The analyzed data from the medial prefrontal cortex had the following quality parameters: signal-to-noise ratio for Glu (M = 13.70, SD = 2.56) and NAA (M = 35.92, SD = 10.14), %CRLB for Glu (M = 2.76, SD = 0.53) and NAA (M = 2.14, SD = 0.58), and FWHM for Glu (M = 26.38, SD = 3.39) and NAA (M = 13.87, SD = 3.50).

The analysis of within-subject effects in the medial prefrontal cortex revealed no significant main effect of stimulation ($F(2, 52) = 0.488$, $p = 0.615$, $\eta^2_p = 0.010$, $BF_{incl} = 0.911$), suggesting that the different types of conditions did not have a major effect on glutamate concentration. However, a statistically significant interaction was found between stimulation and sex ($F(2, 52) = 3.173$, $p = 0.046$, $\eta^2_p = 0.061$, $BF_{incl} = 0.898$). Contrasts revealed differences in females between false font strings and rest ($t(49) = 3.206$, $p = 0.002$, $d = 0.420$), and between words and rest ($t(49) = 2.730$, $p = 0.009$, $d = 0.450$); for both visual stimulations, glutamate concentration was higher than during rest (see Figure 14). No differences between types of stimulation were found in males.

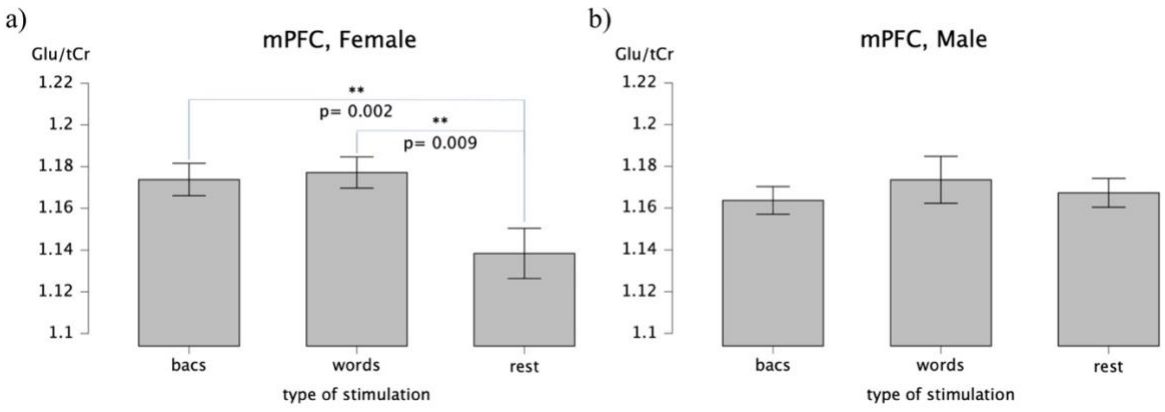

Figure 14. Glutamate concentration changes in response to visual stimulation in the medial prefrontal cortex on 7 Tesla: (a) significant differences in females, with higher glutamate concentration during stimulation than at rest; (b) no significant differences in males. False font strings were presented in BACS font.

No statistically significant effects were observed for the interaction between stimulation and group ($F(2, 52) = 0.597$, $p = 0.553$, $\eta^2_p = 0.012$, $BF_{incl} = 0.171$), between stimulation, group, and sex ($F(2, 52) = 2.008$, $p = 0.140$, $\eta^2_p = 0.039$, $BF_{incl} = 0.621$) or between group and sex ($F(1, 53) = 0.349$, $p = 0.558$, $\eta^2_p = 0.007$, $BF_{incl} = 0.553$). Furthermore, no

significant main effects of group ($F(1, 53) = 1.700$, $p = 0.198$, $\eta^2_p = 0.034$, $BF_{incl} = 0.671$) or sex ($F(1, 53) = 2.842 \cdot 10^{-6}$, $p = 0.999$, $\eta^2_p = 5.799 \cdot 10^{-8}$, $BF_{incl} = 0.400$) were found.

After BOLD correction

No main effect of stimulation was observed ($F(2, 52) = 0.424$, $p = 0.655$, $\eta^2_p = 0.009$, $BF_{incl} = 0.386$). Furthermore, the statistically significant interaction between stimulation and sex was no longer significant after the broadening procedure ($F(2, 52) = 1.796$, $p = 0.171$, $\eta^2_p = 0.035$, $BF_{incl} = 0.300$). Moreover, no statistically significant interaction between stimulation and group was found ($F(2, 52) = 1.090$, $p = 0.340$, $\eta^2_p = 0.022$, $BF_{incl} = 0.299$), between stimulation, group, and sex ($F(2, 52) = 1.664$, $p = 0.195$, $\eta^2_p = 0.033$, $BF_{incl} = 0.515$), between group and sex ($F(1, 53) = 0.205$, $p = 0.653$, $\eta^2_p = 0.004$, $BF_{incl} = 0.518$). Additionally, no significant effect of group ($F(1, 53) = 1.433$, $p = 0.237$, $\eta^2_p = 0.028$, $BF_{incl} = 0.590$) or sex ($F(1, 53) = 0.005$, $p = 0.944$, $\eta^2_p = 1.020 \cdot 10^{-4}$, $BF_{incl} = 0.385$).

Left superior temporal sulcus

The analyzed data from the left superior temporal sulcus had the following quality parameters: signal-to-noise ratio for Glu ($M = 10.79$, $SD = 3.01$) and NAA ($M = 38.74$, $SD = 11.44$), %CRLB for Glu ($M = 3.77$, $SD = 0.96$) and NAA ($M = 2.31$, $SD = 0.76$), and FWHM for Glu ($M = 26.34$, $SD = 2.80$) and NAA ($M = 13.47$, $SD = 2.82$).

In the left superior temporal sulcus, the main effect of stimulation was not significant ($F(2, 46) = 0.257$, $p = 0.774$, $\eta^2_p = 0.006$, $BF_{incl} = 0.163$). However, a statistically significant interaction between stimulation and sex was observed ($F(2, 46) = 4.447$, $p = 0.015$, $\eta^2_p = 0.094$, $BF_{incl} = 1.587$). This effect was driven by higher glutamate concentration in females after visual stimulation with false font strings ($t(43) = 2.436$, $p = 0.019$, $d = 0.307$) compared to rest. Moreover, a trend-level difference was observed in females between words and rest ($t(43) = 1.907$, $p = 0.063$, $d = 0.245$) with higher glutamate concentration after stimulation, and in males between words and rest ($t(43) = -1.884$, $p = 0.066$, $d = -0.247$) with lower glutamate concentration after stimulation (see Figure 15).

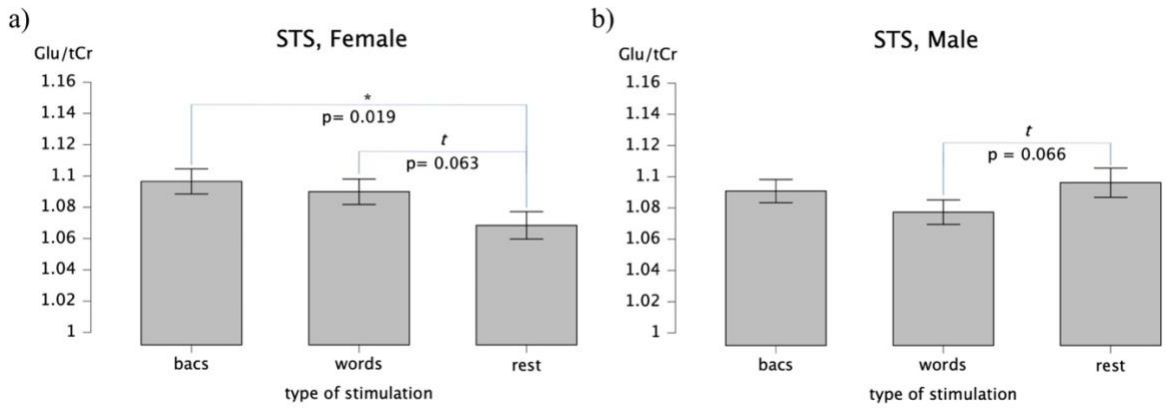

Figure 15. Glutamate concentration changes in response to visual stimulation in the left superior temporal sulcus on 7 Tesla scanner: (a) differences in females; (b) differences in males.

Glutamate concentration also depended on gray matter volume within the analyzed voxels ($F(1, 47) = 8.585$, $p = 0.005$, $\eta^2_p = 0.166$, $BF_{incl} = 12.341$). Additionally, a trend-level effect was observed due to participants' age ($F(1, 47) = 3.258$, $p = 0.078$, $\eta^2_p = 0.070$, $BF_{incl} = 1.692$). No statistically significant interaction between stimulation and group was found ($F(2, 46) = 0.288$, $p = 0.751$, $\eta^2_p = 0.007$, $BF_{incl} = 0.130$), between stimulation, group, and sex ($F(2, 46) = 0.800$, $p = 0.453$, $\eta^2_p = 0.018$, $BF_{incl} = 0.345$), or between group and sex ($F(1, 47) = 0.042$, $p = 0.838$, $\eta^2_p = 9.862 \cdot 10^{-4}$, $BF_{incl} = 0.604$). Additionally, no significant effect of group ($F(1, 47) = 1.757$, $p = 0.192$, $\eta^2_p = 0.039$, $BF_{incl} = 0.833$) or sex ($F(1, 47) = 0.075$, $p = 0.785$, $\eta^2_p = 0.002$, $BF_{incl} = 0.503$).

After BOLD correction

In the left superior temporal sulcus, the main effect of stimulation was not significant ($F(2, 46) = 1.894$, $p = 0.157$, $\eta^2_p = 0.042$, $BF_{incl} = 0.077$). Neither was the interaction between stimulation and sex ($F(2, 46) = 0.375$, $p = 0.689$, $\eta^2_p = 0.009$, $BF_{incl} = 0.290$). A statistically significant effect of gray matter volume was observed ($F(1, 47) = 7.750$, $p = 0.008$, $\eta^2_p = 0.153$, $BF_{incl} = 9.246$), and a trend-level influence of age was revealed ($F(1, 47) = 3.617$, $p = 0.064$, $\eta^2_p = 0.078$, $BF_{incl} = 2.221$). Furthermore, no statistically significant interaction between stimulation and group was revealed ($F(2, 46) = 1.786$, $p = 0.174$, $\eta^2_p = 0.040$, $BF_{incl} = 0.531$), between stimulation, group, and sex ($F(2, 46) = 0.056$, $p = 0.946$, $\eta^2_p = 0.001$, $BF_{incl} = 0.173$), or between group and sex ($F(1, 47) = 0.008$, $p = 0.931$, $\eta^2_p = 1.777 \cdot 10^{-4}$, $BF_{incl} = 0.564$). Additionally, no significant

effect of group ($F(1, 47) = 2.056$, $p = 0.159$, $\eta^2_p = 0.046$, $BF_{incl} = 0.875$) or sex ($F(1, 47) = 0.116$, $p = 0.735$, $\eta^2_p = 0.003$, $BF_{incl} = 0.492$).

3.1.3.3. Data grouped by delay

Medial prefrontal cortex

The analyzed data from the medial prefrontal cortex had the following quality parameters: signal-to-noise ratio for Glu (M = 8.03, SD = 1.45) and NAA (M = 21.00, SD = 5.79), %CRLB for Glu (M = 4.01, SD = 0.92) and NAA (M = 3.12, SD = 0.95), and FWHM for Glu (M = 26.33, SD = 3.35) and NAA (M = 13.87, SD = 3.49).

Results of the ANOVA analysis with visual stimulation type and delays as within-subject factors showed a statistically significant interaction between delay, group, and sex ($F(3, 51) = 2.754$, $p = 0.045$, $\eta^2_p = 0.053$, $BF_{incl} = 1.901$). Contrasts revealed a statistically significant difference for typical reading females between delays of 3000 ms and 4500 ms ($t(49) = 2.042$, $p = 0.047$, $d = 0.336$), with a stronger glutamate response after 3000 ms. Females diagnosed with dyslexia had significantly higher glutamate concentrations 500 ms after stimulation compared to 3000 ms ($t(49) = 2.560$, $p = 0.014$, $d = 0.351$) and 4500 ms ($t(49) = 2.282$, $p = 0.027$, $d = 0.469$). A trend-level difference was observed between 1000 ms (higher glutamate concentration) and 4500 ms ($t(49) = 1.895$, $p = 0.064$, $d = 0.316$). For males with dyslexia, a significantly higher glutamate level was observed 3000 ms after stimulation compared to 1000 ms ($t(49) = -2.914$, $p = 0.005$, $d = -0.500$) and 4500 ms ($t(49) = 2.816$, $p = 0.007$, $d = 0.484$). Additionally, a trend-level difference was found between delays of 500 ms and 1000 ms ($t(49) = 1.742$, $p = 0.088$, $d = 0.307$), with lower glutamate after 1000 ms. A trend-level group difference for males was found after 1000 ms ($t(49) = 1.774$, $p = 0.082$, $d = 0.651$) and 4500 ms ($t(49) = 1.897$, $p = 0.064$, $d = 0.683$), with higher glutamate concentration in the control group compared to the dyslexic group (see Figure 16). No differences were found between the groups of female participants, nor between sexes for both typical readers and those diagnosed with dyslexia.

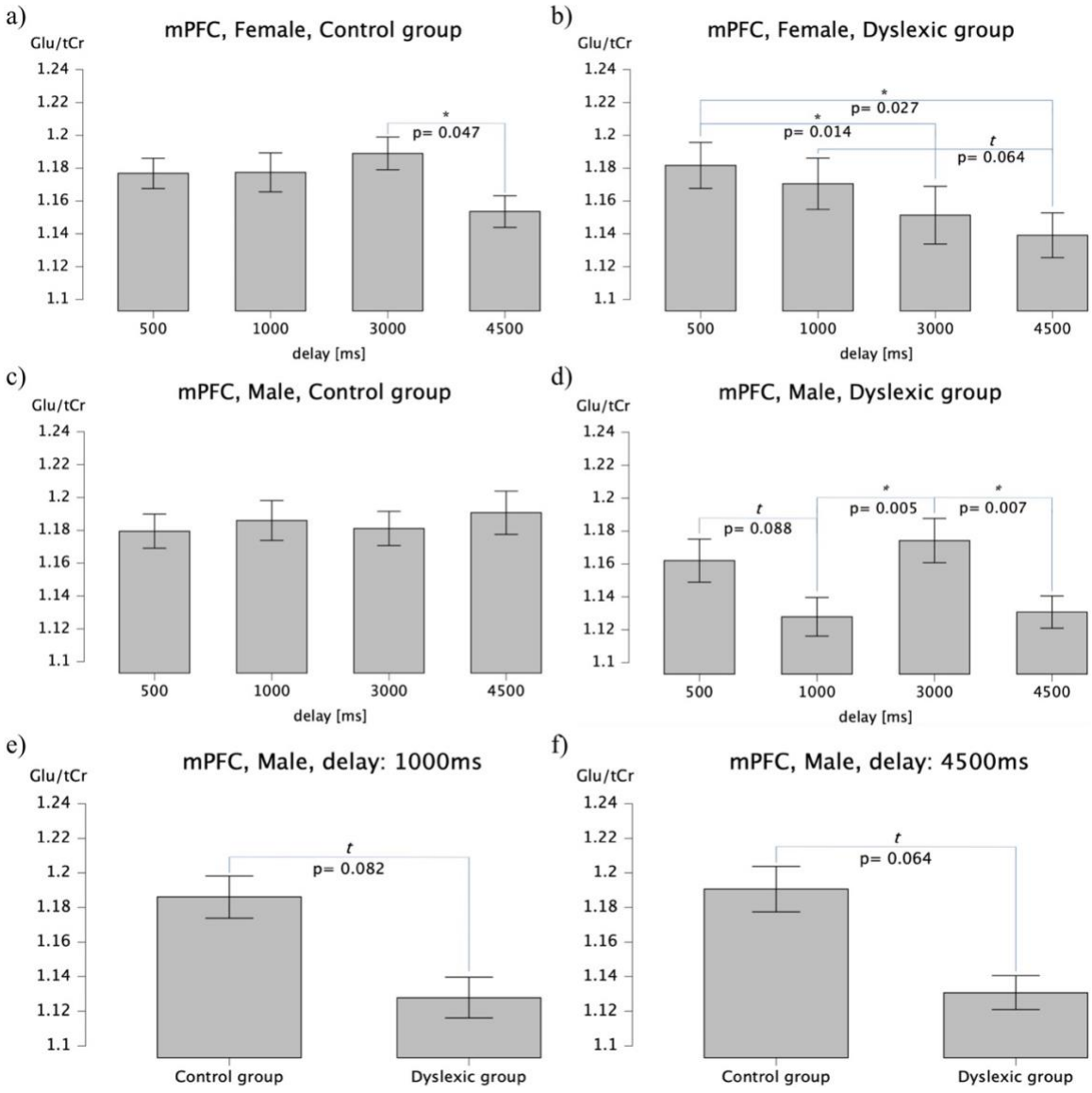

Figure 16. Statistically significant interaction in the medial prefrontal cortex between delay, group, and sex on 7 Tesla scanner: (a) glutamate concentration differences between delays in females from the control group; (b) glutamate concentration differences between delays in females from the dyslexic group; (c) no statistically significant differences between delays in males from the control group; (d) glutamate concentration differences between delays in males from the dyslexic group; (e) glutamate concentration differences between the control and dyslexic groups in males 1000 ms after stimulation; (f) glutamate concentration differences between the control and dyslexic groups in males 4500 ms after stimulation.

No statistically significant effects of delay ($F(3, 51) = 0.142$, $p = 0.934$, $\eta^2_p = 0.003$, $BF_{incl} = 0.329$) and stimulation type ($F(1, 53) = 1.352$, $p = 0.251$, $\eta^2_p = 0.027$, $BF_{incl} = 1.032$) were observed. Moreover no statistically significant interactions between stimulation type and group ($F(1, 53) = 1.504$, $p = 0.226$, $\eta^2_p = 0.030$, $BF_{incl} = 0.314$), between stimulation type and sex ($F(1, 53) = 0.178$, $p = 0.675$, $\eta^2_p = 0.004$, $BF_{incl} = 0.246$),

between stimulation type, group and sex ($F(1, 53) = 0.234$, $p = 0.631$, $\eta^2_p = 0.005$, $BF_{incl} = 0.064$), between delay and group ($F(3, 51) = 1.395$, $p = 0.247$, $\eta^2_p = 0.028$, $BF_{incl} = 0.114$), between delay and sex ($F(3, 51) = 1.763$, $p = 0.157$, $\eta^2_p = 0.035$, $BF_{incl} = 0.248$), between stimulation type and delay ($F(3, 51) = 1.468$, $p = 0.226$, $\eta^2_p = 0.029$, $BF_{incl} = 1.002$), between stimulation type, delay and group ($F(3, 51) = 1.500$, $p = 0.217$, $\eta^2_p = 0.030$, $BF_{incl} = 0.721$), between stimulation type, delay and sex ($F(3, 51) = 0.277$, $p = 0.842$, $\eta^2_p = 0.006$, $BF_{incl} = 0.092$), between stimulation type, delay, group and sex ($F(3, 51) = 0.705$, $p = 0.550$, $\eta^2_p = 0.014$, $BF_{incl} = 0.029$) and between group and sex ($F(1, 53) = 0.281$, $p = 0.599$, $\eta^2_p = 0.006$, $BF_{incl} = 0.418$) were observed. In addition, no statistically significant effects of group ($F(1, 53) = 1.824$, $p = 0.183$, $\eta^2_p = 0.036$, $BF_{incl} = 0.737$) and sex ($F(1, 53) = 0.076$, $p = 0.783$, $\eta^2_p = 0.002$, $BF_{incl} = 0.342$) were found.

After BOLD correction

After the BOLD correction procedure, the statistically significant interaction between delay, group, and sex was no longer observed. Additionally, no other effects or interactions were found to be significant or at trend-level. No statistically significant effects of delay ($F(3, 51) = 0.847$, $p = 0.470$, $\eta^2_p = 0.017$, $BF_{incl} = 0.136$) and stimulation type ($F(1, 53) = 0.260$, $p = 0.612$, $\eta^2_p = 0.005$, $BF_{incl} = 0.423$) were revealed. Furthermore no statistically significant interactions between stimulation type and group ($F(1, 53) = 0.659$, $p = 0.421$, $\eta^2_p = 0.013$, $BF_{incl} = 0.301$), between stimulation type and sex ($F(1, 53) = 0.493$, $p = 0.486$, $\eta^2_p = 0.010$, $BF_{incl} = 0.234$), between stimulation type, group and sex ($F(1, 53) = 0.007$, $p = 0.936$, $\eta^2_p = 1.334 \cdot 10^{-4}$, $BF_{incl} = 0.292$), between delay and group ($F(3, 51) = 1.045$, $p = 0.375$, $\eta^2_p = 0.021$, $BF_{incl} = 0.073$), between delay and sex ($F(3, 51) = 1.911$, $p = 0.130$, $\eta^2_p = 0.038$, $BF_{incl} = 0.174$), between delay, group and sex ($F(3, 51) = 0.961$, $p = 0.413$, $\eta^2_p = 0.019$, $BF_{incl} = 0.057$), between stimulation type and delay ($F(3, 51) = 1.866$, $p = 0.138$, $\eta^2_p = 0.037$, $BF_{incl} = 0.645$), between stimulation type, delay and group ($F(3, 51) = 0.789$, $p = 0.502$, $\eta^2_p = 0.016$, $BF_{incl} = 0.171$), between stimulation type, delay and sex ($F(3, 51) = 0.468$, $p = 0.705$, $\eta^2_p = 0.009$, $BF_{incl} = 0.414$), between stimulation type, delay, group and sex ($F(3, 51) = 0.734$, $p = 0.533$, $\eta^2_p = 0.015$, $BF_{incl} = 0.027$) and between group and sex ($F(1, 53) = 0.247$, $p = 0.622$, $\eta^2_p = 0.005$, $BF_{incl} = 0.494$) were observed. In addition, no statistically significant effects of group

(F(1, 53) = 1.554, p = 0.218, η^2_p = 0.031, BF_{incl} = 0.560) and sex (F(1, 53) = 0.042, p = 0.839, $\eta^2_p = 8.485 \cdot 10^{-4}$, BF_{incl} = 0.350) were found.

Left superior temporal sulcus

The analyzed data from the left superior temporal sulcus had the following quality parameters: signal-to-noise ratio for Glu (M = 6.28, SD = 1.73) and NAA (M = 22.53, SD = 6.59), %CRLB for Glu (M = 5.23, SD = 1.49) and NAA (M = 3.18, SD = 1.22), and FWHM for Glu (M = 26.32, SD = 2.78) and NAA (M = 13.42, SD = 2.80).

The main effects of delays and stimulation type were not statistically significant (F(3, 45) = 1.604, p = 0.192, η^2_p = 0.036, BF_{incl} = 0.054), (F(3, 45) = 0.651, p = 0.424, η^2_p = 0.015, BF_{incl} = 0.176) respectively. No statistically significant interactions between stimulation type and group (F(1, 53) = 0.028, p = 0.869, $\eta^2_p = 6.427 \cdot 10^{-4}$, BF_{incl} = 0.183), between stimulation type and sex (F(1, 53) = 0.083, p = 0.774, η^2_p = 0.002, BF_{incl} = 0.372), between stimulation type, group and sex (F(1, 53) = $1.759 \cdot 10^{-4}$, p = 0.989, $\eta^2_p = 4.091 \cdot 10^{-6}$, BF_{incl} = 0.602), between delay and group (F(3, 51) = 0.450, p = 0.718, η^2_p = 0.010, BF_{incl} = 0.017), between delay and sex (F(3, 51) = 1.437, p = 0.235, η^2_p = 0.032, BF_{incl} = 0.097), between delay, group and sex (F(3, 51) = 0.571, p = 0.635, η^2_p = 0.013, BF_{incl} = 0.077), between stimulation type, delay and group (F(3, 51) = 0.680, p = 0.566, η^2_p = 0.016, BF_{incl} = 0.243), between stimulation type, delay and sex (F(3, 51) = 1.148, p = 0.332, η^2_p = 0.026, BF_{incl} = 4.840), between stimulation type, delay, group and sex (F(3, 51) = 0.744, p = 0.528, η^2_p = 0.017, BF_{incl} = 0.067) and between group and sex (F(1, 53) = 0.178, p = 0.675, η^2_p = 0.004, BF_{incl} = 0.604) were observed. In addition, no statistically significant effects of group (F(1, 53) = 1.740, p = 0.194, η^2_p = 0.039, BF_{incl} = 1.334) and sex (F(1, 53) = 0.028, p = 0.867, $\eta^2_p = 6.553 \cdot 10^{-4}$, BF_{incl} = 0.328) were found.

Interestingly, a trend-level interaction between stimulation type and delay was observed (F(3, 45) = 2.272, p = 0.083, η^2_p = 0.050, BF_{incl} = 0.082). This was due to an increase in glutamate concentration 4500 ms after word stimulation compared to 500 ms (t(43) = -1.749, p = 0.087, d = -0.228) (see Figure 17).

Glutamate concentration in the left superior temporal sulcus also depended on gray matter volume within the analyzed voxels ($F(1, 47) = 7.643$, $p = 0.008$, $\eta^2_p = 0.151$, $BF_{incl} = 5.550$) and age ($F(1, 47) = 4.759$, $p = 0.035$, $\eta^2_p = 0.100$, $BF_{incl} = 2.204$).

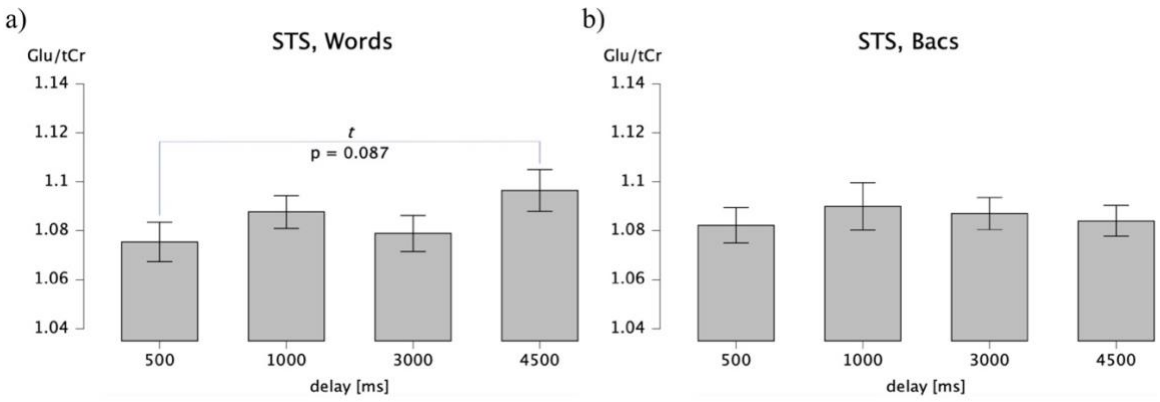

Figure 17. Glutamate concentration at different time delays in the left superior temporal sulcus on 7 Tesla scanner: (a) trend-level difference after word stimulation; (b) no statistically significant differences after false font string stimulation.

After BOLD correction

After the BOLD correction procedure, a trend-level interaction between stimulation type and delay remained significant ($F(3, 45) = 2.317$, $p = 0.079$, $\eta^2_p = 0.051$, $BF_{incl} = 0.040$); however, contrasts did not reveal the underlying cause for this. Additionally, gray matter volume ($F(1, 47) = 7.646$, $p = 0.008$, $\eta^2_p = 0.151$, $BF_{incl} = 7.529$) and age ($F(1, 47) = 4.712$, $p = 0.036$, $\eta^2_p = 0.099$, $BF_{incl} = 2.450$) significantly influenced glutamate concentration.

No statistically significant effects of delay ($F(3, 45) = 0.449$, $p = 0.719$, $\eta^2_p = 0.010$, $BF_{incl} = 0.048$) and stimulation type ($F(1, 53) = 0.191$, $p = 0.664$, $\eta^2_p = 0.004$, $BF_{incl} = 0.109$) were revealed. No statistically significant interactions between stimulation type and group ($F(1, 53) = 1.198$, $p = 0.280$, $\eta^2_p = 0.027$, $BF_{incl} = 0.265$), between stimulation type and sex ($F(1, 53) = 1.257$, $p = 0.268$, $\eta^2_p = 0.028$, $BF_{incl} = 0.257$), between stimulation type, group and sex ($F(1, 53) = 0.004$, $p = 0.947$, $\eta^2_p = 1.037 \cdot 10^{-4}$, $BF_{incl} = 0.450$), between delay and group ($F(3, 51) = 0.063$, $p = 0.979$, $\eta^2_p = 0.001$, $BF_{incl} = 0.015$), between delay and sex ($F(3, 51) = 1.443$, $p = 0.233$, $\eta^2_p = 0.032$, $BF_{incl} = 0.063$), between delay, group and sex ($F(3, 51) = 0.241$, $p = 0.867$, $\eta^2_p = 0.006$, $BF_{incl} = 0.052$), between stimulation type, delay and group ($F(3, 51) = 0.045$, $p = 0.987$,

η^2_p = 0.001, BF_{incl} = 0.008), between stimulation type, delay and sex (F(3, 51) = 0.972, p = 0.408, η^2_p = 0.022, BF_{incl} = 0.014), between stimulation type, delay, group and sex (F(3, 51) = 0.641, p = 0.590, η^2_p = 0.015, BF_{incl} = 0.071) and between group and sex (F(1, 53) = 0.171, p = 0.681, η^2_p = 0.004, BF_{incl} = 0.585) were observed. Additionally, no statistically significant effects of group (F(1, 53) = 1.892, p = 0.176, η^2_p = 0.042, BF_{incl} = 1.209) and sex (F(1, 53) = 0.054, p = 0.818, η^2_p = 0.001, BF_{incl} = 0.370) were found.

3.1.3.4. Data analyzed using the dynamic-averaged approach

Medial prefrontal cortex

The model, considering both words and false font strings, with data analyzed using the dynamic module in FSL-MRS, did not reveal a significant main effect of delay (F(3, 51) = 1.586, p = 0.195, η^2_p = 0.031, BF_{incl} = 0.623) or stimulation type (F(1, 53) = 2.415, p = 0.127, η^2_p = 0.047, BF_{incl} = 0.574). Furthermore, no statistically significant interactions between stimulation type and group (F(1, 53) = 0.248, p = 0.620, η^2_p = 0.005, BF_{incl} = 0.228), between stimulation type and sex (F(1, 53) = 1.139, p = 0.291, η^2_p = 0.023, BF_{incl} = 0.188), between stimulation type, group and sex (F(1, 53) = 0.168, p = 0.684, η^2_p = 0.003, BF_{incl} = 0.260), between delay and group (F(3, 51) = 0.541, p = 0.655, η^2_p = 0.011, BF_{incl} = 0.058), between delay and sex (F(3, 51) = 1.466, p = 0.226, η^2_p = 0.029, BF_{incl} = 0.103), between delay, group and sex (F(3, 51) = 1.658, p = 0.179, η^2_p = 0.033, BF_{incl} = 0.364), between stimulation type and delay (F(3, 51) = 1.943, p = 0.125, η^2_p = 0.038, BF_{incl} = 0.084), between stimulation type, delay and sex (F(3, 51) = 0.188, p = 0.904, η^2_p = 0.004, BF_{incl} = 0.033), between stimulation type, delay, group and sex (F(3, 51) = 0.335, p = 0.800, η^2_p = 0.007, BF_{incl} = 0.014) and between group and sex (F(1, 53) = 0.534, p = 0.468, η^2_p = 0.011, BF_{incl} = 0.737) were observed. Moreover, no statistically significant effects of group (F(1, 53) = 1.447, p = 0.235, η^2_p = 0.029, BF_{incl} = 0.647) and sex (F(1, 53) = 0.005, p = 0.943, $\eta^2_p = 1.042 \cdot 10^{-4}$, BF_{incl} = 0.481) were found.

However, a significant interaction between stimulation type, delay, and group was observed (F(3, 51) = 3.089, p = 0.029, η^2_p = 0.059, BF_{incl} = 7.021) (see Figure 18). This effect was driven by differences between words (lower glutamate levels) and false font strings at 1000 ms after stimulation in the dyslexic group (t(49) = -2.746, p = 0.008, d = -0.420), and

between words (lower glutamate concentration) and false font strings at 3000 ms delay in the control group ($t(49) = -2.538$, $p = 0.014$, $d = -0.305$). Additionally, differences were observed between 500 ms and 1000 ms ($t(49) = 2.955$, $p = 0.005$, $d = 0.441$) and 1000 ms and 3000 ms ($t(49) = -2.792$, $p = 0.007$, $d = -0.443$) for individuals with dyslexia stimulated with words (weaker glutamate response after 1000 ms). Moreover, differences between 500 ms and 3000 ms ($t(49) = -2.049$, $p = 0.046$, $d = -0.232$), 1000 ms and 3000 ms ($t(49) = -2.641$, $p = 0.011$, $d = -0.308$), and 3000 ms and 4500 ms ($t(49) = 2.098$, $p = 0.041$, $d = 0.232$) were observed for the control group stimulated using false font strings, with a stronger glutamate response after 3000 ms. A trend-level difference was found between 500 ms and 4500 ms (lower glutamate) after word stimulation in the control group ($t(49) = 1.989$, $p = 0.052$, $d = 0.272$). Furthermore, a group difference between typical readers and individuals diagnosed with dyslexia was observed 1000 ms after word stimulation ($t(49) = 2.038$, $p = 0.047$, $d = 0.648$), with higher glutamate concentration in the control group.

A trend-level interaction was observed between delay and gray matter volume ($F(3, 51) = 2.481$, $p = 0.063$, $\eta^2_p = 0.048$).

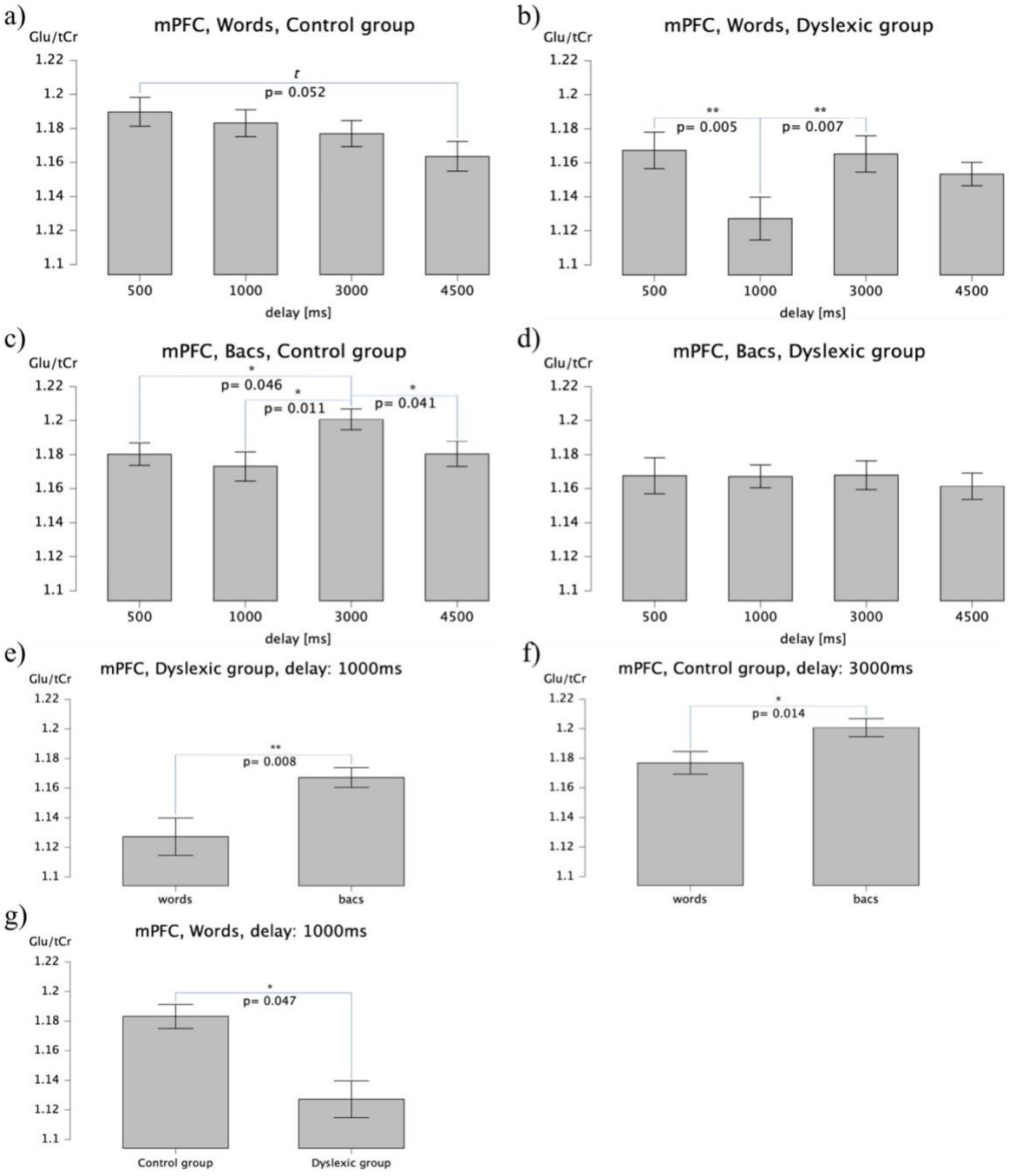

Figure 18. Statistically significant differences in glutamate concentration in the medial prefrontal cortex using the dynamic-averaged approach on 7 Tesla scanner: (a) between delays in the control group after word stimulation; (b) between delays in the dyslexic group after word stimulation; (c) between delays in the control group after false font string stimulation; (d) no statistically significant difference between delays in the dyslexic group after false font string stimulation; (e) between word and false font string stimulation in the dyslexic group at 1000 ms; (f) between word and false font string stimulation in the control group at 3000 ms; (g) between control and dyslexic groups at 1000 ms after word stimulation.

After BOLD correction

Data analyzed after the bordering procedure showed no significant effects of delay ($F(3, 51) = 1.742$, $p = 0.161$, $\eta^2_p = 0.034$, $BF_{incl} = 0.262$) or stimulation type ($F(1, 53) = 0.140$, $p = 0.710$, $\eta^2_p = 0.003$, $BF_{incl} = 0.276$). No statistically significant interactions between stimulation type and group ($F(1, 53) = 1.333$, $p = 0.254$, $\eta^2_p = 0.026$, $BF_{incl} = 0.478$), between stimulation type and sex ($F(1, 53) = 2.327$, $p = 0.134$, $\eta^2_p = 0.045$, $BF_{incl} = 0.482$), between stimulation type, group and sex ($F(1, 53) = 1.093 \cdot 10^{-6}$, $p = 0.999$, $\eta^2_p = 2.231 \cdot 10^{-8}$, $BF_{incl} = 0.205$), between delay and group ($F(3, 51) = 0.272$, $p = 0.846$, $\eta^2_p = 0.006$, $BF_{incl} = 0.052$), between delay and sex ($F(3, 51) = 1.357$, $p = 0.258$, $\eta^2_p = 0.027$, $BF_{incl} = 0.111$), between stimulation type and delay ($F(3, 51) = 1.583$, $p = 0.196$, $\eta^2_p = 0.031$, $BF_{incl} = 0.553$), between stimulation type, delay and sex ($F(3, 51) = 0.095$, $p = 0.963$, $\eta^2_p = 0.002$, $BF_{incl} = 0.036$), between stimulation type, delay, group and sex ($F(3, 51) = 0.399$, $p = 0.754$, $\eta^2_p = 0.008$, $BF_{incl} = 1.462$) and between group and sex ($F(1, 53) = 0.776$, $p = 0.383$, $\eta^2_p = 0.016$, $BF_{incl} = 0.610$) were observed. Additionally, no statistically significant effects of group ($F(1, 53) = 1.559$, $p = 0.218$, $\eta^2_p = 0.031$, $BF_{incl} = 0.628$) and sex ($F(1, 53) = 0.011$, $p = 0.916$, $\eta^2_p = 2.287 \cdot 10^{-4}$, $BF_{incl} = 0.474$) were found.

However, a statistically significant interaction between delay and gray matter volume was revealed ($F(3, 51) = 2.835$, $p = 0.040$, $\eta^2_p = 0.055$,). A trend-level interaction between delay, group, and sex ($F(3, 51) = 2.211$, $p = 0.089$, $\eta^2_p = 0.043$, $BF_{incl} = 0.442$) was driven by differences in typical reading females between 500 ms and 4500 ms ($t(49) = 2.166$, $p = 0.035$, $d = 0.269$) and between 3000 ms and 4500 ms (trend-level difference: $t(49) = 1.936$, $p = 0.059$, $d = 0.250$), with a weaker glutamate response after 4500 ms. Additionally, for females with dyslexia, a trend-level difference was found between 500 ms (higher glutamate) and 3000 ms ($t(49) = 1.870$, $p = 0.067$, $d = 0.237$). A difference between delays of 1000 ms and 3000 ms ($t(49) = -2.713$, $p = 0.009$, $d = -0.401$), and between 3000 ms and 4500 ms ($t(49) = 2.281$, $p = 0.027$, $d = 0.308$), with higher glutamate concentration at 3000 ms after stimulation, was observed in males diagnosed with dyslexia. Additionally, in this group a trend-level difference was found between delays of 500 ms (higher glutamate) and 1000 ms ($t(49) = 1.692$, $p = 0.097$, $d = 0.222$). Moreover, a trend-level difference was found for males 1000 ms after stimulation between the control

group (higher glutamate) and the dyslexia group (t(49) = 1.697, p = 0.096, d = 0.633) (see Figure 19).

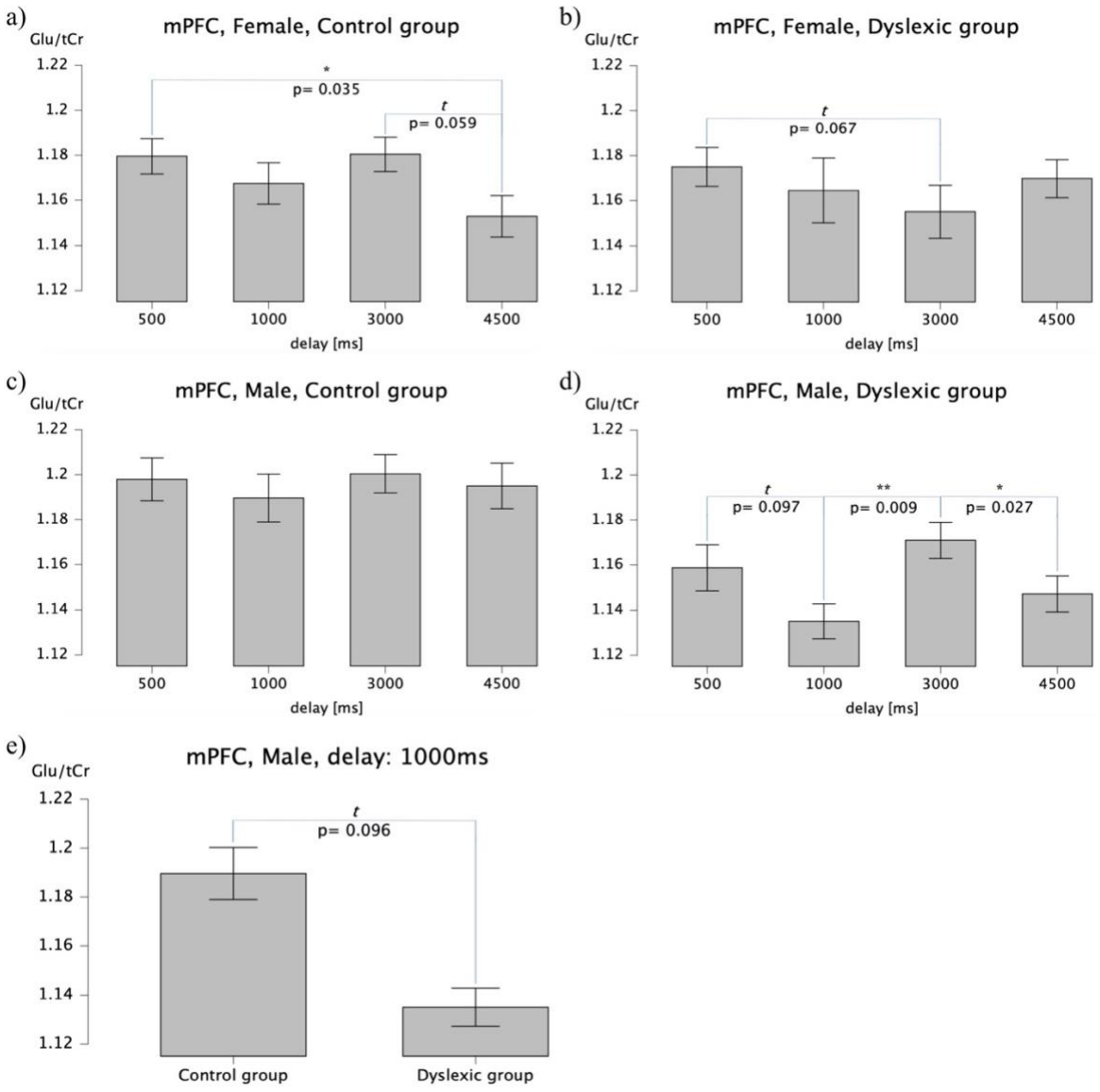

Figure 19. Difference in glutamate concentration in the medial prefrontal cortex after BOLD correction on 7 Tesla scanner showing interaction between delay, group, and sex: (a) differences in females from the control group; (b) differences in females from the dyslexic group; (c) no statistically significant differences in males from the control group; (d) differences in males from the dyslexic group; € difference between control and dyslexic groups in males at 1000 ms after visual stimulation.

Another trend-level interaction between stimulation type, delay, and group ($F(3, 51) = 2.268$, $p = 0.083$, $\eta^2_p = 0.044$, $BF_{incl} = 2.401$) was a result of a lower glutamate level 1000 ms after word and stimulation in the dyslexia group compared to 500 ms ($t(49) = 2.750$, $p = 0.008$, $d = 0.393$), 3000 ms ($t(49) = -2.073$, $p = 0.043$, $d = -0.280$), and

at a trend-level difference to 4500 ms ($t(49) = -1.775$, $p = 0.082$, $d = -0.287$). Additionally, in the typical reading group glutamate was higher 500 ms after word stimulation than after 4500 ms ($t(49) = 2.069$, $p = 0.044$, $d = 0.282$). In the control group, a stronger glutamate response 3000 ms after stimulation with false font strings was observed compared to 500 ms ($t(49) = -2.112$, $p = 0.040$, $d = -0.211$), 1000 ms ($t(49) = -2.222$, $p = 0.031$, $d = -0.279$), and 4500 ms ($t(49) = 1.997$, $p = 0.051$, $d = 0.221$), as a trend-level difference. Furthermore, 1000 ms after word stimulation, a higher glutamate concentration in the control group compared to the dyslexia group was found ($t(49) = 2.027$, $p = 0.048$, $d = 0.616$). In individuals with dyslexia, 1000 ms after word stimulation, glutamate levels were lower than after false font strings ($t(49) = -2.625$, $p = 0.012$, $d = -0.408$). In typical readers, a similar trend-level observation was found 3000 ms after stimulation, with a weaker glutamate response after words ($t(49) = -1.876$, $p = 0.067$, $d = -0.232$) (see Figure 20).

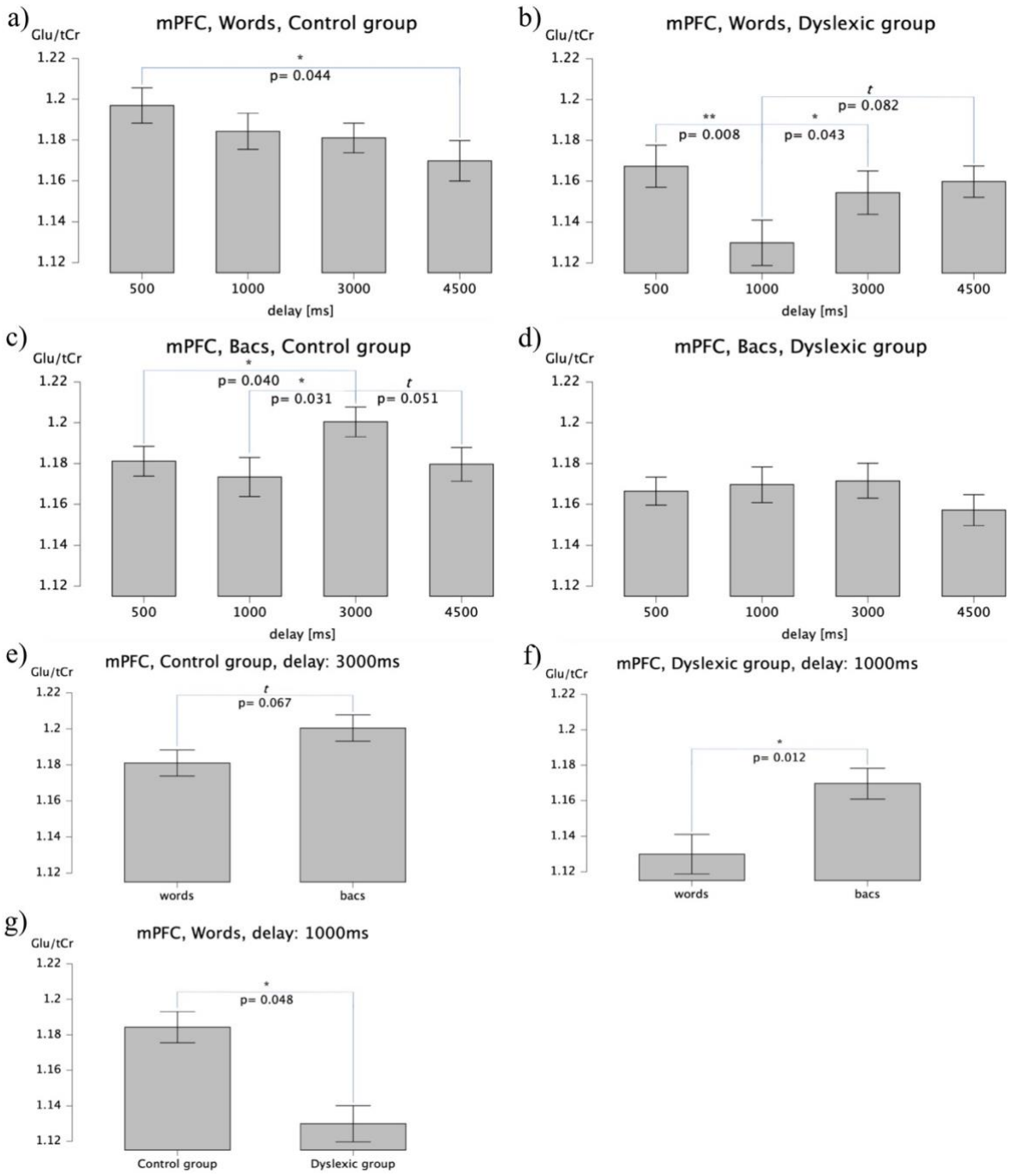

Figure 20. Trend-level interaction between stimulation type, delay, and group in the medial prefrontal cortex after BOLD correction on 7 Tesla scanner: (a) differences in glutamate concentration between delays in the control group after word stimulation; (b) differences in the dyslexic group after word stimulation; (c) differences in the control group after false font string stimulation; (d) no statistically significant differences in the dyslexic group after false font string stimulation; € differences in glutamate concentration after word and false font string stimulation at 3000 ms in the control group; (f) differences after word and false font string stimulation at 1000 ms in the dyslexic group; (g) differences between the control and dyslexic groups at 1000 ms after word stimulation.

Left superior temporal sulcus

In the left superior temporal sulcus, the dynamic-averaged approach revealed significant effects of gray matter volume ($F(1, 47) = 8.121$, $p = 0.007$, $\eta^2_p = 0.159$, $BF_{incl} = 8.558$) and age ($F(1, 47) = 5.251$, $p = 0.027$, $\eta^2_p = 0.109$, $BF_{incl} = 3.662$). The main effect of delay was not statistically significant ($F(3, 45) = 0.357$, $p = 0.784$, $\eta^2_p = 0.008$, $BF_{incl} = 0.032$), also no statistically significant effects of stimulation type was revealed ($F(1, 53) = 1.844$, $p = 0.182$, $\eta^2_p = 0.041$, $BF_{incl} = 0.613$). Furthermore, no statistically significant interactions between stimulation type and group ($F(1, 53) = 1.075$, $p = 0.306$, $\eta^2_p = 0.024$, $BF_{incl} = 0.364$), between stimulation type and sex ($F(1, 53) = 0.041$, $p = 0.841$, $\eta^2_p = 9.474 \cdot 10^{-4}$, $BF_{incl} = 0.243$), between stimulation type, group and sex ($F(1, 53) = 0.191$, $p = 0.665$, $\eta^2_p = 0.004$, $BF_{incl} = 0.091$), between delay and group ($F(3, 51) = 0.820$, $p = 0.485$, $\eta^2_p = 0.019$, $BF_{incl} = 0.046$), between delay and sex ($F(3, 51) = 1.495$, $p = 0.219$, $\eta^2_p = 0.034$, $BF_{incl} = 0.151$), between delay, group and sex ($F(3, 51) = 1.312$, $p = 0273$, $\eta^2_p = 0.030$, $BF_{incl} = 0.414$), between stimulation type and delay ($F(3, 51) = 1.092$, $p = 0.355$, $\eta^2_p = 0.025$, $BF_{incl} = 0.057$), between stimulation type, delay and group ($F(3, 51) = 0.663$, $p = 0.576$, $\eta^2_p = 0.015$, $BF_{incl} = 0.105$), between stimulation type, delay and sex ($F(3, 51) = 1.496$, $p = 0.219$, $\eta^2_p = 0.034$, $BF_{incl} = 0.112$), between stimulation type, delay, group and sex ($F(3, 51) = 0.591$, $p = 0.622$, $\eta^2_p = 0.014$, $BF_{incl} = 0.133$) and between group and sex ($F(1, 53) = 0.079$, $p = 0.780$, $\eta^2_p = 0.002$, $BF_{incl} = 0.477$) were observed. No statistically significant effects of group ($F(1, 53) = 1.834$, $p = 0.183$, $\eta^2_p = 0.041$, $BF_{incl} = 0.975$) and sex ($F(1, 53) = 0.140$, $p = 0.710$, $\eta^2_p = 0.003$, $BF_{incl} = 0.555$) were found.

After BOLD Correction

The broadening procedure revealed a trend-level significant four-way interaction between stimulation type, delay, group, and sex ($F(3, 45) = 2.283$, $p = 0.082$, $\eta^2_p = 0.050$, $BF_{incl} = 0.163$). Contrasts show that this interaction was due to differences between delays in females with dyslexia stimulated with false font strings, where glutamate levels were lower at 3000 ms compared to 500 ms ($t(43) = 3.492$, $p = 0.001$, $d = 0.600$), and 4500 ms ($t(43) = -2.022$, $p = 0.049$, $d = -0.367$), and at 1000 ms at trend-level ($t(43) = 1.890$, $p = 0.065$, $d = 0.372$). Additionally, differences between delays were observed for males diagnosed with dyslexia after false font strings, with lower glutamate concentration at

500 ms compared to 3000 ms ($t(43) = -2.536$, $p = 0.015$, $d = -0.507$), and higher glutamate concentration at 3000 ms compared to 4500 ms ($t(43) = 2.914$, $p = 0.006$, $d = 0.615$), along with a trend-level difference between 1000 ms (higher glutamate concentration) and 4500 ms ($t(43) = 1.784$, $p = 0.082$, $d = 0.353$). Moreover, dyslexic females showed a stronger glutamate response at 3000 ms after word stimulation compared to false font strings ($t(43) = 2.225$, $p = 0.031$, $d = 0.493$). Notably, dyslexic males had a trend-level weaker glutamate response at 3000 ms after word stimulation compared to false font strings ($t(43) = -1.862$, $p = 0.070$, $d = -0.480$). A trend-level difference was also observed in individuals with dyslexia at 3000 ms after false font string stimulation, with an difference between females and males ($t(43) = -1.770$, $p = 0.084$, $d = -0.687$), with higher glutamate levels in males. A trend-level glutamate difference between the control and dyslexic groups was found in females at 3000 ms after word stimulation ($t(43) = -1.922$, $p = 0.061$, $d = -0.796$) and at 500 ms after false font strings ($t(43) = -1.847$, $p = 0.072$, $d = -0.735$), with higher glutamate concentration in individuals with dyslexia in both cases (see Figure 21).

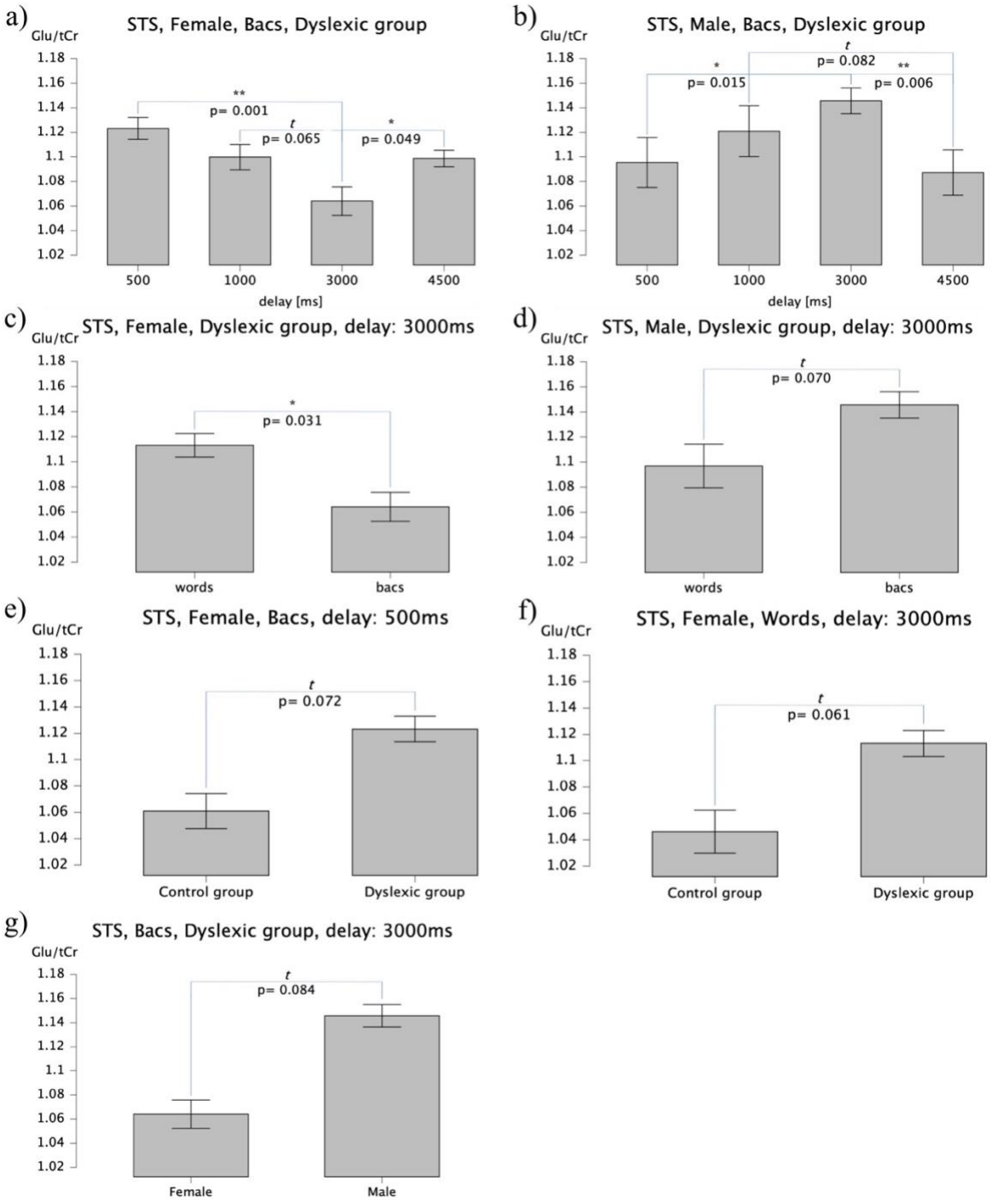

Figure 21. Differences in glutamate concentration in the left superior temporal sulcus after BOLD correction on 7 Tesla scanner showing a four-way interaction between stimulation type, delay, group, and sex: (a) differences between delays in females from the dyslexic group after false font string stimulation; (b) differences between delays in males from the dyslexic group after false font string stimulation; (c) difference between word and false font string stimulation in females from the dyslexic group at 3000 ms; (d) difference between word and false font string stimulation in males from the dyslexic group at 3000 ms; (e) difference between control and dyslexic groups in females at 500 ms after false font string stimulation; (f) difference between control and dyslexic groups in females at 3000 ms after word stimulation;

(g) difference between females and males from the dyslexic group at 3000 ms after false font string stimulation.

Age ($F(1, 47) = 6.420$, $p = 0.015$, $\eta^2_p = 0.130$, $BF_{incl} = 5.315$) and gray matter volume ($F(1, 47) = 7.939$, $p = 0.007$, $\eta^2_p = 0.156$, $BF_{incl} = 9.769$) were statistically significant in the dynamic-averaged approach for the left superior temporal sulcus.

Analysis did not reveal a significant main effect of delay ($F(3, 45) = 0.107$, $p = 0.956$, $\eta^2_p = 0.002$, $BF_{incl} = 0.028$) or stimulation type ($F(1, 53) = 0.193$, $p = 0.662$, $\eta^2_p = 0.004$, $BF_{incl} = 0.442$). Moreover, no statistically significant interactions between stimulation type and group ($F(1, 53) = 1.594$, $p = 0.214$, $\eta^2_p = 0.036$, $BF_{incl} = 0.351$), between stimulation type and sex ($F(1, 53) = 0.060$, $p = 0.807$, $\eta^2_p = 0.001$, $BF_{incl} = 0.258$), between stimulation type, group and sex ($F(1, 53) = 1.077$, $p = 0.305$, $\eta^2_p = 0.024$, $BF_{incl} = 0.351$), between delay and group ($F(3, 51) = 0.687$, $p = 0.562$, $\eta^2_p = 0.016$, $BF_{incl} = 0.049$), between delay and sex ($F(3, 51) = 1.071$, $p = 0.364$, $\eta^2_p = 0.024$, $BF_{incl} = 0.026$), between delay, group and sex ($F(3, 51) = 0.973$, $p = 0.408$, $\eta^2_p = 0.022$, $BF_{incl} = 0.050$), between stimulation type and delay ($F(3, 51) = 0.242$, $p = 0.867$, $\eta^2_p = 0.006$, $BF_{incl} = 0.210$), between stimulation type, delay and group ($F(3, 51) = 0.208$, $p = 0.891$, $\eta^2_p = 0.005$, $BF_{incl} = 0.678$), between stimulation type, delay and sex ($F(3, 51) = 1.727$, $p = 0.165$, $\eta^2_p = 0.039$, $BF_{incl} = 0.569$) and between group and sex ($F(1, 53) = 0.349$, $p = 0.558$ $\eta^2_p = 0.008$, $BF_{incl} = 0.546$) were observed. In addition, no statistically significant effects of group ($F(1, 53) = 1.857$, $p = 0.180$, $\eta^2_p = 0.041$, $BF_{incl} = 0.929$) and sex ($F(1, 53) = 0.051$, $p = 0.822$, $\eta^2_p = 0.001$, $BF_{incl} = 0.395$) were found.

3.2. Experiment 2 – fMRS on 3T scanner

3.2.1. Behavioral results

Functional magnetic resonance spectroscopy (fMRS) was performed on a 3 Tesla scanner, involving 40 participants—21 diagnosed with dyslexia and 19 typical readers. The participants were matched for sex, age, and family socio-economic status, as indicated by the years of education of their parents.

Volunteers diagnosed with dyslexia scored higher on the Polish version of the Adult Reading History Questionnaire (ARHQ-PL), suggesting an increased risk for dyslexia. Participants with dyslexia exhibited lower scores in intelligence quotients (IQ), although both groups remained within the normal intellectual range. A trend-level difference in nonverbal IQ was observed, with typical readers showing higher scores. No significant effects of sex or interactions between sex and group were found (see Table 6).

Reading ability differed between participants from the control and dyslexic groups. Reading speed for both words and pseudowords was significantly higher in the control group. Typical readers also performed better in phonological awareness tasks. Interestingly, while the dyslexic group generally responded slower in rapid automatized naming tasks, statistically significant group differences were found for colors and letters, whereas only a trend-level difference was observed for objects and digits. Moreover, a statistically significant difference between sexes was found in rapid automatized naming of colors, with female participants responding faster, although this effect was no longer significant after correction for multiple comparisons. No other sex differences or interactions between sex and group were observed (see Table 7).

Table 6. Descriptive Statistics for Demographic Characteristics Separately for Females and Males in Dyslexic and Control Groups. For all Comparisons, F Statistics, P-Values (in Brackets), and the Partial Eta Squared (η^2_p) are Provided

	DYS F (n = 9)		CON F (n = 11)		DYS M (n = 12)		CON M (n = 8)		group $F(1,38)$	η^2_p	sex $F(1,38)$	η^2_p	group*sex $F(1, 38)$	η^2_p
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*						
Age	22.653	2.771	21.907	3.359	20.641	3.675	21.826	2.431	0.047 (0.830)	0.001	1.056 (0.311)	0.028	0.899 (0.349)	0.024
Mother's education (years)	15.444	2.963	15.727	2.494	17.542	3.986	16.875	2.232	0.038 (0.847)	0.001	2.701 (0.109)	0.070	0.231 (0.634)	0.006
Father's education (years)	15.222	3.866	16.818	4.262	17.333	2.309	17.000	3.854	0.302 (0.586)	0.008	0.994 (0.325)	0.027	0.704 (0.407)	0.019
IQ	105.000	14.782	112.545	7.866	103.917	13.201	116.750	8.225	7.659 (0.009)	0.175	0.180 (0.674)	0.005	0.516 (0.477)	0.014
Nonverbal IQ (scaled score)	10.667	3.240	11.909	2.468	10.167	3.486	12.500	2.138	3.612 (0.065)	0.091	0.002 (0.962)	6.484×10^{-5}	0.336 (0.566)	0.009
ARHQ-PL	50.667	10.488	24.000	5.079	48.083	11.389	23.375	7.425	78.471 (< .001)	0.686	0.306 (0.584)	0.008	0.114 (0.738)	0.003

Note. CON – control group; DYS – dyslexic group; F – females; M – males; ARHQ-PL – Polish version of the Adult Reading History Questionnaire.
Boldface indicates statistical significance at $p < .05$ level.

Table 7. Behavioral results from reading and reading-related tasks. For all comparisons, F-statistics, p-values (in brackets), and the partial eta squared (η^2_p) are provided.

	DYS F (n = 9)		CON F (n = 11)		DYS M (n = 12)		CON M (n = 8)		group $F(1,38)$	η^2_p	sex $F(1,38)$	η^2_p	group*sex $F(1,38)$	η^2_p
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*						
words/min	115.889	17.808	137.364	10.726	113.667	17.541	130.625	12.872	15.742 (< .001)*	0.304	0.856 (0.361)	0.023	0.217 (0.644)	0.006
pseudowords/min	59.667	15.116	84.727	15.193	60.833	15.724	87.500	19.086	24.957 (< .001)*	0.409	0.145 (0.706)	0.004	0.024 (0.878)	6.678×10^{-4}
reading comprehension (sec)	58.333	9.695	41.636	6.577	61.917	18.158	46.750	10.430	15.983 (< .001)*	0.307	1.191 (0.282)	0.032	0.037 (0.849)	0.001
RAN objects (sec)	31.333	5.852	28.818	5.828	33.667	6.679	29.875	2.997	2.990 (0.092)	0.077	0.864 (0.359)	0.023	0.122 (0.728)	0.003
RAN colors (sec)	33.556	5.659	28.818	3.545	37.917	6.960	31.250	2.659	11.832 (0.001)*	0.247	4.198 (0.048)	0.104	0.339 (0.564)	0.009
RAN digits (sec)	18.889	2.759	17.182	3.157	20.000	6.281	16.625	2.825	3.484 (0.070)	0.088	0.041 (0.840)	0.001	0.375 (0.544)	0.010
RAN letters (sec)	22.444	4.503	18.818	1.168	23.250	4.827	20.750	2.375	6.982 (0.012)	0.162	1.394 (0.245)	0.037	0.236 (0.630)	0.007
phoneme deletion (% correct)	77.777	29.214	94.055	5.796	75.641	30.741	97.115	3.409	7.088 (0.012)	0.165	0.004 (0.948)	1.179×10^{-4}	0.134 (0.716)	0.004
spoonerisms phonemes (% correct)	73.809	29.014	88.312	13.277	45.833	38.939	88.391	9.304	11.079 (0.002)*	0.235	2.648 (0.112)	0.069	2.678 (0.110)	0.069

spoonerisms syllables (% correct)	59.259	30.174	71.212	27.978	41.666	30.567	81.248	13.908	8.712 (0.006)	0.195	0.187 (0.668)	0.005	2.504 (0.122)	0.065
orthographic awareness (accuracy/time)	0.377	0.125	0.622	0.112	0.350	0.131	0.589	0.119	38.251 (< .001)*	0.515	0.583 (0.450)	0.016	0.007 (0.935)	1.859×10^{-4}

Note. CON – control group; DYS – dyslexic group; F – females; M – males; RAN – rapid automatized naming. Boldface indicates statistical significance at $p < .05$ level.
* Significance after Bonferroni correction for 11 planned comparisons for reading-related tasks; threshold for statistical significance after correction was set at p = 0.00455;

3.2.2. Task During fMRS Scanning

The tasks during the fMRS scan on the 3T scanner were exactly the same as those on the 7T scanner. Participants were instructed to press the response pad whenever a target word appeared on the screen. The percentage of correct responses was determined by comparing the number of correct clicks to the total number of target words presented during the scanning task.

Consistent with the 7T scanner task, only scans with spectra recognized as good quality were included in the analysis.

A similar number of total responses (clicks on the pad) was observed for both analyzed regions, with the mean number for the medial prefrontal cortex: 107.289 (SD = 11.680, Min = 74, Max = 138), and for the left superior temporal sulcus: 104.543 (SD = 13.267, Min = 61, Max = 141).

The medial prefrontal cortex was scanned first 15 times, second 16 times, and third 7 times. The left superior temporal sulcus was scanned at the beginning of the fMRS session on the 3T scanner 22 times, and as the second region 13 times.

Univariate ANOVA on percentage accuracy revealed a statistically significant difference between groups in both brain regions, with typical readers showing better accuracy than readers with dyslexia. Moreover, a statistically significant interaction was found between group and sex in the medial prefrontal cortex (see Table 8), driven by higher accuracy in typical reading males compared to those diagnosed with dyslexia ($t(34) = 3.010$, $p = 0.005$, $d = 1.455$), and a difference between females and males in the dyslexic group, with females showing higher accuracy than males ($t(34) = 2.317$, $p = 0.027$, $d = 1.042$). For the brain region involved in reading (STS), a trend-level interaction between group and sex was observed (see Table 3), driven by a group difference among males, with the control males group showing better scores than dyslexic males ($t(31) = 2.702$, $p = 0.011$, $d = 1.307$). Additionally, a difference in accuracy between females (better accuracy) and males in the dyslexic group was found ($t(31) = 2.291$, $p = 0.029$, $d = 1.108$).

Table 8. Behavioral results from the task performed during the scan for data that meet the quality criteria. For all comparisons, F-statistics, p-values (in brackets), and the partial eta squared (η^2_p) are provided.

Target recognition (% correct)	DYS F		CON F		DYS M		CON M		group *F*(1,36) *F*(1,33)	η^2_p	sex *F*(1,36) *F*(1,33)	η^2_p	group*sex *F*(1,36) *F*(1,33)	η^2_p
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*						
Medial prefrontal cortex	71.154	5.065	71.678	4.351	63.549	11.843	74.176	1.612	5.353 (0.027)	0.136	1.123 (0.297)	0.032	4.393 (0.044)	0.114
Left superior temporal sulcus	70.467	3.409	71.635	4.824	59.003	17.331	72.527	2.535	4.227 (0.048)	0.120	2.188 (0.149)	0.066	2.990 (0.094)	0.088

3.2.3. fMRS Results

3.2.3.1. Quality Assessment

To ensure the quality of the data in the packages was sufficient for analysis, the signal-to-noise ratio (SNR) and the percentage of Cramer-Rao lower bound (%CRLB) for two metabolites, glutamate and NAA, were reported by the FSL-MRS software. Additionally, the full width at half maximum (FWHM) was checked, although only for NAA due to software issues in accurately assessing FWHM for glutamate. All packages met the quality criteria: %CRLB for glutamate should be less than 20% (the maximum %CRLB for glutamate in the package with signals acquired 4500 ms after word stimulation was 13.610 in the medial prefrontal cortex, and in the left superior temporal sulcus for the same type of signals, words 4500 ms, it was 13.622). Data quality details can be found in Table 9 and Table 10 for FWHM.

Table 9. Quality parameters of the data, grouped based on stimulation type and delay for 3 Tesla scanner.

Packages type	Max # signals	Glutamate				NAA			
		Mean	SD	Min	Max	Mean	SD	Min	Max
Signal-to-noise ratio (SNR) in mPFC (38 participants)									
Words	78	10.398	2.664	5.285	16.768	33.212	9.186	15.573	52.119
Bacs	78	10.361	2.550	5.408	15.271	33.020	9.029	15.752	51.069
Rest	78	10.201	2.488	5.125	15.381	32.492	8.767	16.090	52.260
words 500 ms	26	6.038	1.621	2.503	9.970	18.985	5.165	8.846	29.681
words 1000 ms	26	5.991	1.523	2.938	9.344	19.454	5.363	9.296	30.647
words 3000 ms	26	5.985	1.427	3.039	9.235	19.049	5.050	9.206	29.374
words 4500 ms	26	6.001	1.407	3.051	8.803	18.915	5.038	9.159	28.781
bacs 500 ms	26	6.049	1.535	3.211	9.105	19.261	5.340	8.970	31.802
bacs 1000 ms	26	6.090	1.539	2.671	9.147	19.292	5.323	8.457	30.816
bacs 3000 ms	26	5.903	1.464	3.222	8.827	19.016	5.117	9.667	28.175
bacs 4500 ms	26	5.966	1.462	3.055	9.985	18.916	5.095	9.065	30.062
Signal-to-noise ratio (SNR) is STS (35 participants)									
Words	78	10.288	2.625	5.533	15.234	39.584	7.553	20.875	57.415
Bacs	78	10.094	2.566	4.872	14.735	39.217	7.153	21.750	56.474
Rest	78	10.068	2.538	5.096	14.443	39.304	6.698	21.954	53.302
words 500 ms	26	5.933	1.457	3.183	9.298	22.893	4.069	12.164	32.147
words 1000 ms	26	5.826	1.494	2.724	8.739	22.527	4.251	11.816	32.462
words 3000 ms	26	5.930	1.477	3.280	8.318	22.817	4.100	12.634	33.678
words 4500 ms	26	5.874	1.450	2.683	8.520	22.768	4.216	12.102	34.172

bacs 500 ms	26	5.815	1.534	2.983	8.890	22.561	4.106	12.846	33.567
bacs 1000 ms	26	5.829	1.558	2.657	8.924	22.578	4.086	11.947	31.609
bacs 3000 ms	26	5.855	1.501	2.763	8.554	22.694	4.240	12.450	34.598
bacs 4500	26	5.840	1.385	3.463	8.873	22.633	3.790	13.479	31.413
%CRLB in mPFC (38 participants)									
Words	78	3.829	0.875	2.783	7.313	2.629	0.694	1.583	4.994
Bacs	78	3.813	0.780	2.649	5.866	2.679	0.692	1.484	4.713
Rest	78	3.849	0.855	2.813	6.712	2.694	0.769	1.451	5.119
words 500 ms	26	5.920	1.958	3.569	14.282	4.199	1.378	2.050	8.249
words 1000 ms	26	5.803	1.531	3.821	10.557	3.903	1.357	2.222	8.053
words 3000 ms	26	5.784	1.557	3.878	11.204	4.000	1.133	2.311	8.011
words 4500 ms	26	5.772	1.752	3.846	13.610	3.973	1.209	2.249	8.453
bacs 500 ms	26	5.759	1.408	3.668	9.000	3.970	1.085	2.107	6.725
bacs 1000 ms	26	5.720	1.598	3.811	10.628	3.955	1.405	2.035	8.341
bacs 3000 ms	26	5.851	1.376	3.927	9.169	4.008	1.217	2.051	6.186
bacs 4500 ms	26	5.808	1.544	3.696	10.705	3.949	1.141	2.204	6.790
%CRLB in STS (35 participants)									
Words	78	4.798	1.504	3.100	9.791	2.443	0.701	1.457	4.085
Bacs	78	4.899	1.583	3.150	9.037	2.463	0.672	1.573	3.858
Rest	78	4.850	1.502	3.234	8.815	2.473	0.663	1.579	3.959
words 500 ms	26	6.665	1.851	4.128	11.649	3.443	0.769	1.861	5.211
words 1000 ms	26	6.823	2.086	4.673	13.281	3.573	0.931	2.248	6.596
words 3000 ms	26	6.634	1.998	4.289	12.696	3.336	0.920	1.988	5.980
words 4500 ms	26	6.710	2.128	4.502	13.622	3.382	0.811	1.973	5.135
bacs 500 ms	26	6.854	2.150	4.114	13.215	3.438	0.899	2.062	6.384
bacs 1000 ms	26	6.774	2.163	4.414	13.267	3.367	0.929	2.273	6.245
bacs 3000 ms	26	6.804	2.223	4.369	12.611	3.407	0.897	2.037	6.140
bacs 4500 ms	26	6.631	1.692	4.314	10.729	3.398	0.693	2.019	4.969

Table 10. Full width at half maximum for NAA of the data, grouped based on stimulation type and delay for 3 Tesla scanner.

Packages type	Max # signals	**mPFC (38 participants)**				**STS (35 participants)**			
		Mean	SD	Min	Max	Mean	SD	Min	Max
FWHM for NAA									
Words	78	4.761	0.702	3.703	7.023	5.510	0.965	4.381	8.555
Bacs	78	4.791	0.675	3.702	7.050	5.551	1.043	4.297	8.707
Rest	78	4.776	0.684	3.600	6.933	5.508	0.938	4.261	8.264
words 500 ms	26	4.777	0.714	3.709	7.245	5.513	0.960	4.399	8.562
words 1000 ms	26	4.739	0.670	3.583	6.533	5.546	1.010	4.281	8.641
words 3000 ms	26	4.765	0.737	3.554	7.218	5.492	0.937	4.423	8.483
words 4500 ms	26	4.778	0.705	3.663	6.633	5.522	0.989	4.185	8.759
bacs 500 ms	26	4.812	0.705	3.651	6.900	5.535	1.063	4.337	8.734
bacs 1000 ms	26	4.742	0.701	3.791	7.247	5.570	0.970	4.289	8.564
bacs 3000 ms	26	4.791	0.648	3.602	6.256	5.535	1.095	4.219	8.815
bacs 4500 ms	26	4.774	0.726	3.694	7.260	5.543	0.936	4.441	8.879

3.2.3.2. Data Exclusion Due to BOLD Correction Issues

Due to technical issues, the apodization procedure, which is used to broaden spectra peaks in order to reduce potential influences of the BOLD effect on spectra shape, was not properly applied to some data packages. As a result, one spectrum from the medial prefrontal cortex was omitted from the analysis for both stimulation type and delay-based grouping, as well as for the dynamic-averaged approach, since the same data packages were used in both analyses. In the left superior temporal sulcus, data from two participants were excluded from the stimulation type analysis, and data from three participants were omitted from the analysis of delays between stimulation and data acquisition for both the traditional and dynamic-averaged approaches.

Finally, analysis was conducted on data from 37 participants for the medial prefrontal cortex: 18 participants from the control group (11 females, 7 males) and 19 from the dyslexic group (9 females, 10 males). For the left superior temporal sulcus, 33 participants' data were used for analysis grouped by stimulation type: 17 participants from the control group (10 females, 7 males) and 16 from the dyslexic group (6 females, 10 males). Additionally, 32 participants' data were analyzed based on delays and using the dynamic-averaged approach: 17 participants from the control group (10 females, 7 males) and 15 from the dyslexic group (5 females, 10 males).

3.2.3.3. Data grouped by stimulation type

Medial prefrontal cortex

The analyzed data from the medial prefrontal cortex had the following quality parameters: signal-to-noise ratio for Glu (M = 10.32, SD = 2.57) and NAA (M = 32.91, SD = 8.99), %CRLB for Glu (M = 3.83, SD = 0.84) and NAA (M = 2.67, SD = 0.72), and FWHM for NAA (M = 4.78, SD = 0.69).

In the medial prefrontal cortex, repeated measures ANOVA analysis did not reveal a statistically significant main effect of stimulation ($F(2, 35) = 0.243$, $p = 0.785$, $\eta^2_p = 0.008$, $BF_{incl} = 0.213$). No statistically significant effects were observed for the interaction between stimulation and group ($F(2, 35) = 0.394$, $p = 0.676$, $\eta^2_p = 0.012$, $BF_{incl} = 0.232$), between stimulation and sex ($F(2, 35) = 1.073$, $p = 0.348$, $\eta^2_p = 0.032$, $BF_{incl} = 0.319$), between stimulation, group, and sex ($F(2, 35) = 0.562$, $p = 0.573$, $\eta^2_p = 0.017$, $BF_{incl} = 0.350$) or

between group and sex ($F(1, 36) = 0.288$, $p = 0.595$, $\eta^2_p = 0.009$, $BF_{incl} = 0.526$). Furthermore, no significant effects of group ($F(1, 36) = 0.050$, $p = 0.825$, $\eta^2_p = 0.002$, $BF_{incl} = 0.375$) or sex ($F(1, 36) = 1.281$, $p = 0.266$, $\eta^2_p = 0.039$, $BF_{incl} = 0.746$) were found.

A trend-level effect of age was found ($F(1, 36) = 4.020$, $p = 0.053$, $\eta^2_p = 0.112$, $BF_{incl} = 2.542$).

After BOLD correction

The same data following the apodization procedure showed only a trend-level effect of participants' age ($F(1, 35) = 3.617$, $p = 0.067$, $\eta^2_p = 0.104$, $BF_{incl} = 2.189$). The effect of stimulation type was not statistically significant ($F(2, 34) = 0.800$, $p = 0.454$, $\eta^2_p = 0.025$, $BF_{incl} = 0.169$). No statistically significant effects were observed for the interaction between stimulation and group ($F(2, 34) = 0.317$, $p = 0.729$, $\eta^2_p = 0.010$, $BF_{incl} = 0.205$), between stimulation and sex ($F(2, 34) = 0.420$, $p = 0.659$, $\eta^2_p = 0.013$, $BF_{incl} = 0.222$), between stimulation, group, and sex ($F(2, 34) = 1.253$, $p = 0.293$, $\eta^2_p = 0.039$, $BF_{incl} = 0.639$) or between group and sex ($F(1, 35) = 0.318$, $p = 0.577$, $\eta^2_p = 0.010$, $BF_{incl} = 0.516$). No significant effects of group ($F(1, 35) = 0.057$, $p = 0.814$, $\eta^2_p = 0.002$, $BF_{incl} = 0.348$) or sex ($F(1, 35) = 1.427$, $p = 0.241$, $\eta^2_p = 0.044$, $BF_{incl} = 0.751$) were observed.

Left superior temporal sulcus

The analyzed data from the left superior temporal sulcus had the following quality parameters: signal-to-noise ratio for Glu (M = 10.15, SD = 2.58) and NAA (M = 39.37, SD = 7.13), %CRLB for Glu (M = 4.85, SD = 1.53) and NAA (M = 2.46, SD = 0.68), and FWHM for NAA (M = 5.52, SD = 0.98).

In the superior temporal sulcus, no statistically significant effect of stimulation type was observed ($F(2, 32) = 1.624$, $p = 0.206$, $\eta^2_p = 0.053$, $BF_{incl} = 0.810$). No statistically significant effects were observed for the interaction between stimulation and group ($F(2, 32) = 0.023$, $p = 0.978$, $\eta^2_p = 7.762 \cdot 10^{-4}$, $BF_{incl} = 0.169$), between stimulation and sex ($F(2, 32) = 1.775$, $p = 0.179$, $\eta^2_p = 0.058$, $BF_{incl} = 0.793$), between stimulation, group, and sex ($F(2, 32) = 0.049$, $p = 0.952$, $\eta^2_p = 0.002$, $BF_{incl} = 0.235$) or between group and sex ($F(1, 33) = 0.534$, $p = 0.471$, $\eta^2_p = 0.018$, $BF_{incl} = 0.831$). No significant effects of group

($F(1, 33) = 1.372$, $p = 0.251$, $\eta^2_p = 0.045$, $BF_{incl} = 0.756$) or sex ($F(1, 33) = 0.160$, $p = 0.692$, $\eta^2_p = 0.006$, $BF_{incl} = 0.754$) were revealed.

However, it was found that gray matter volume within the analyzed voxel had a statistically significant effect on glutamate concentration ($F(1, 33) = 7.801$, $p = 0.009$, $\eta^2_p = 0.212$, $BF_{incl} = 3.600$).

After BOLD correction

The statistically significant effect of stimulation ($F(2, 30) = 5.382$, $p = 0.007$, $\eta^2_p = 0.166$, $BF_{incl} = 0.806$) was observed, driven by the difference in glutamate concentration between words (higher glutamate concentration) and the rest ($t(27) = 2.743$, $p = 0.011$, $d = 0.219$) (see Figure 22). Additionally, a statistically significant interaction was observed between stimulation and sex ($F(2, 30) = 3.216$, $p = 0.048$, $\eta^2_p = 0.106$, $BF_{incl} = 1.602$), caused by changes in glutamate levels in females between false font strings and words ($t(27) = -2.322$, $p = 0.028$, $d = -0.360$) and between words and the rest ($t(27) = 3.515$, $p = 0.002$, $d = 0.423$), with higher glutamate levels following word stimulation. A trend-level difference between females and males was also observed for false font strings ($t(27) = -2.033$, $p = 0.052$, $d = -0.733$) and the rest ($t(27) = -1.882$, $p = 0.071$, $d = -0.725$) (see Figure 23). Furthermore, a statistically significant effect of group was identified ($F(1, 31) = 4.784$, $p = 0.038$, $\eta^2_p = 0.151$, $BF_{incl} = 1.036$), with higher glutamate concentrations in the control group than in the dyslexia group (see Figure 24). Moreover, gray matter volume had a statistically significant influence on glutamate levels ($F(1, 31) = 6.363$, $p = 0.018$, $\eta^2_p = 0.191$, $BF_{incl} = 3.151$). In addition, a trend-level interaction appeared between group and sex ($F(1, 31) = 3.471$, $p = 0.073$, $\eta^2_p = 0.114$, $BF_{incl} = 1.301$), determined by differences between females in the typical reading and dyslexic groups, with weaker glutamate responses in the dyslexic group ($t(27) = 2.814$, $p = 0.009$, $d = 1.437$), and a difference between females and males in the dyslexic group ($t(27) = -2.322$, $p = 0.028$, $d = -1.250$), with higher glutamate concentrations in males (see Figure 25).

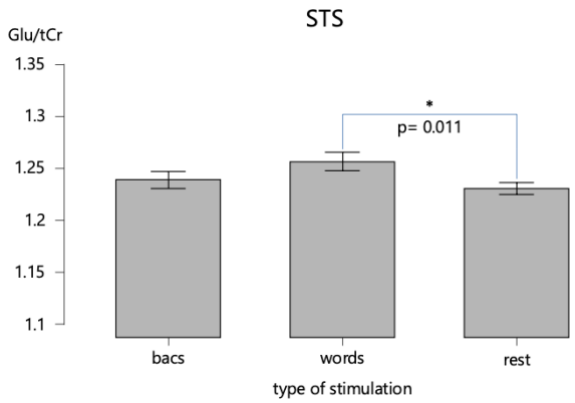

Figure 22. Glutamate concentration changes in response to visual stimulation after BOLD correction in left superior temporal sulcus observed on 3 Tesla scanner.

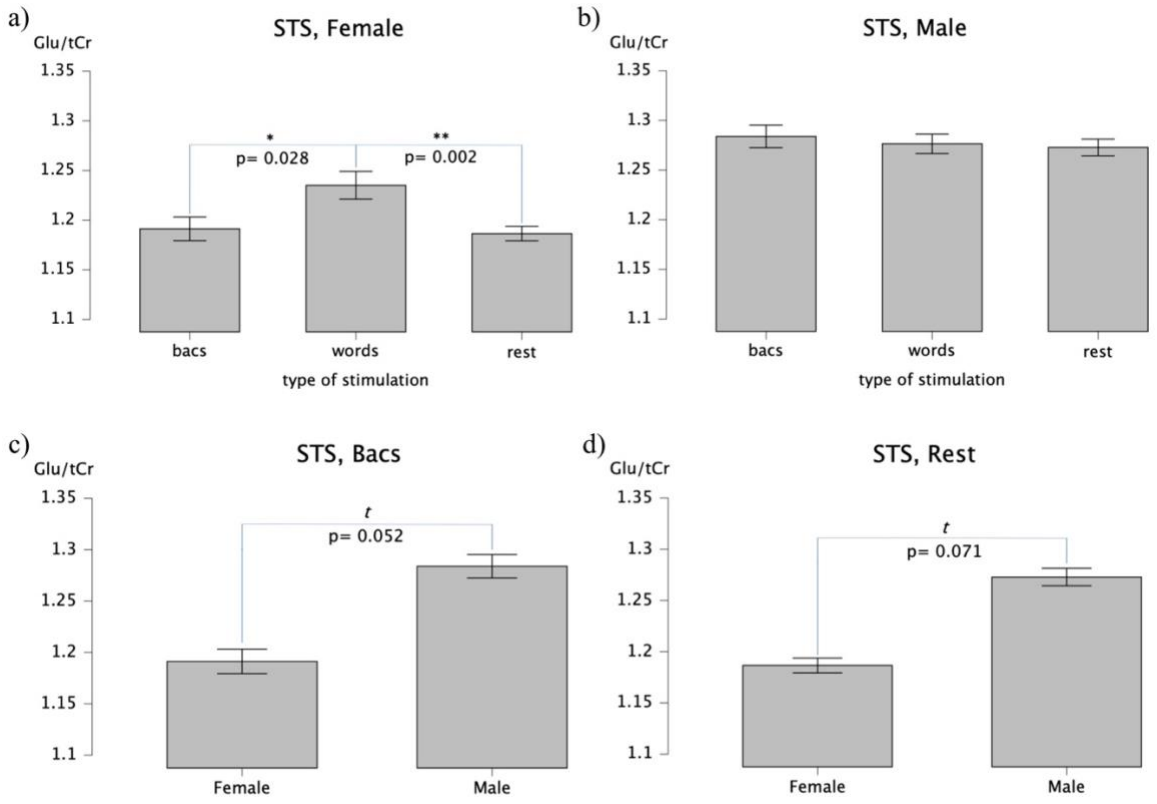

Figure 23. Statistically significant interaction between stimulation and sex on 3 Tesla scanner: a) glutamate concentration changes in response to visual stimulation in females; b) no glutamate concentration changes in response to visual stimulation in males; c) difference between females and males after false font string stimulation (bacs font); d) difference between females and males during rest.

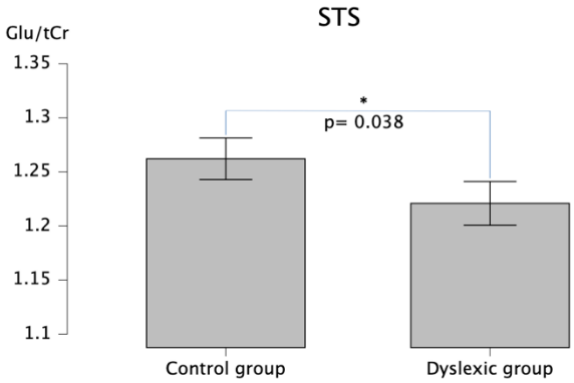

Figure 24. Difference in glutamate concentration between the control and dyslexic groups after BOLD correction in the left superior temporal sulcus observed on a 3 Tesla scanner.

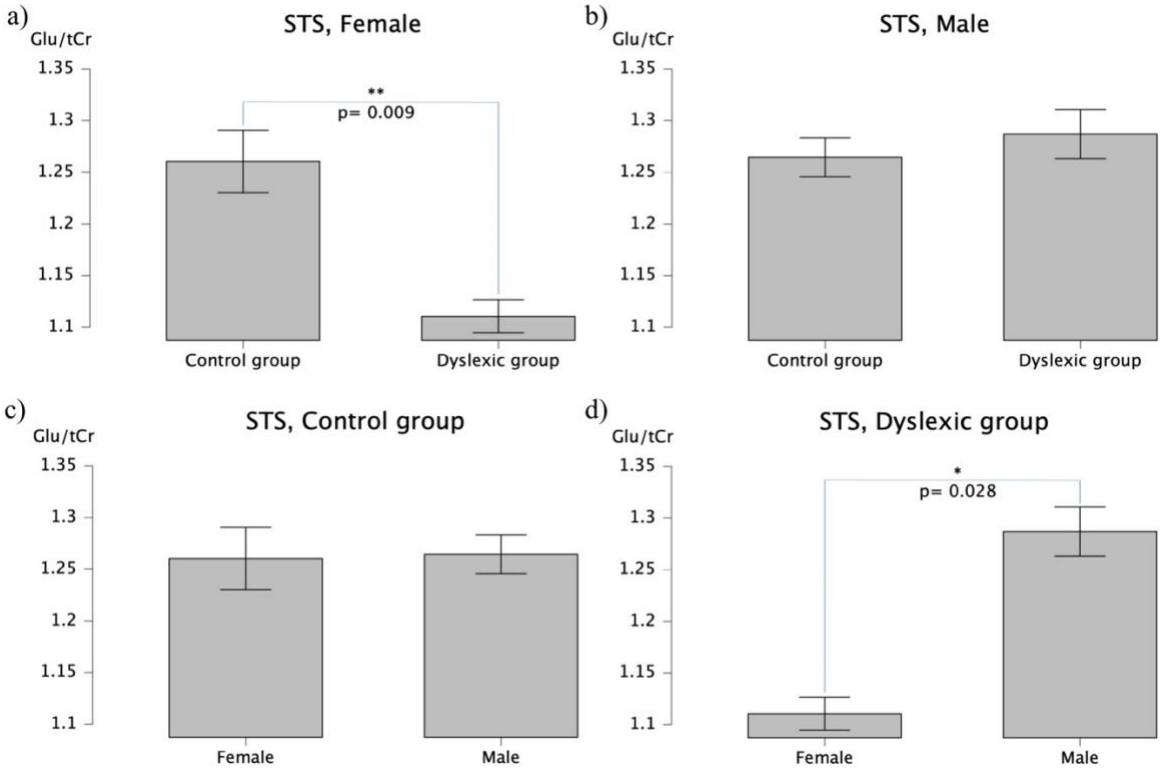

Figure 25. a) Difference in glutamate concentration between the control and dyslexic groups in females; b) no differences between the control and dyslexic groups in males. Males had higher glutamate concentration than females in the dyslexic (d) but not the control group (c) on 3 Tesla scanner.

No statistically significant effects were observed for the interaction between stimulation and group ($F(2, 30) = 0.165$, $p = 0.849$, $\eta^2_p = 0.006$, $BF_{incl} = 0.242$) or between stimulation, group, and sex ($F(2, 30) = 0.008$, $p = 0.992$, $\eta^2_p = 3.079 \cdot 10^{-4}$, $BF_{incl} = 0.223$). No significant effect of sex ($F(1, 31) = 2.638$, $p = 0.116$, $\eta^2_p = 0.089$, $BF_{incl} = 0.974$) was found.

3.2.3.4. Data grouped by delay

Medial prefrontal cortex

The analyzed data from the medial prefrontal cortex had the following quality parameters: signal-to-noise ratio for Glu (M = 6.00, SD = 1.50) and NAA (M = 19.11, SD = 5.19), %CRLB for Glu (M = 5.80, SD = 1.59) and NAA (M = 3.99, SD = 1.24), and FWHM for NAA (M = 4.77, SD = 0.70).

Analysis taking into account different time delays between stimulation onset and signal acquisition in the medial prefrontal cortex discovered a statistically significant three-way interaction between stimulation type, group, and sex ($F(1, 36) = 5.263$, $p = 0.028$, $\eta^2_p = 0.141$, $BF_{incl} = 1.223$), determined by difference in typical reading females between stimulation with words and false font strings ($t(32) = -2.145$, $p = 0.040$, $d = -0.332$), with higher glutamate levels after false font string stimulation (see Figure 26). Moreover, the interaction between stimulation type and delay was significant ($F(3, 34) = 4.038$, $p = 0.009$, $\eta^2_p = 0.112$, $BF_{incl} = 0.134$), caused by a trend-level difference between 500 ms and 3000 ms ($t(32) = 2.029$, $p = 0.051$, $d = 0.408$), and between 3000 ms and 4500 ms ($t(32) = -1.716$, $p = 0.096$, $d = -0.353$) after false font string stimulation, with weaker glutamate response after 3000 ms (see Figure 27). A statistically significant interaction between stimulation type, delay, and age ($F(3, 34) = 3.184$, $p = 0.027$, $\eta^2_p = 0.090$), as well as an interaction between stimulation type, delay, and gray matter volume ($F(3, 34) = 3.173$, $p = 0.028$, $\eta^2_p = 0.090$), was observed. Additionally, a trend-level effect of delay was reported ($F(3, 34) = 2.325$, $p = 0.080$, $\eta^2_p = 0.068$, $BF_{incl} = 0.043$), driven by higher glutamate concentration 4500 ms after stimulation compared to 1000 ms ($t(32) = -1.756$, $p = 0.089$, $d = -0.223$) and 3000 ms ($t(32) = -1.941$, $p = 0.061$, $d = -0.236$) (see Figure 28). A trend-level interaction was also observed between delay and age ($F(3, 34) = 2.263$, $p = 0.086$, $\eta^2_p = 0.066$). A statistically significant effect of age ($F(1, 36) = 4.401$, $p = 0.044$, $\eta^2_p = 0.121$, $BF_{incl} = 5.792$) was revealed.

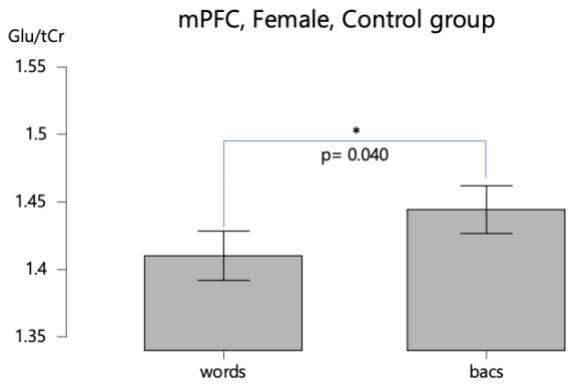

Figure 26. Difference in glutamate concentration on 3 Tesla scanner in females from the control group between stimulation with words and false font strings (written in bacs font).

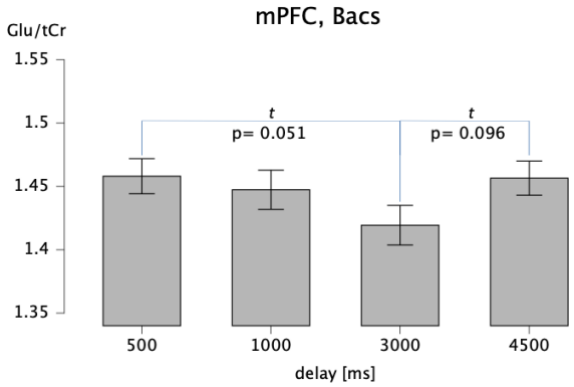

Figure 27. Glutamate concentration differences on 3 Tesla scanner between time delays after false font string stimulation.

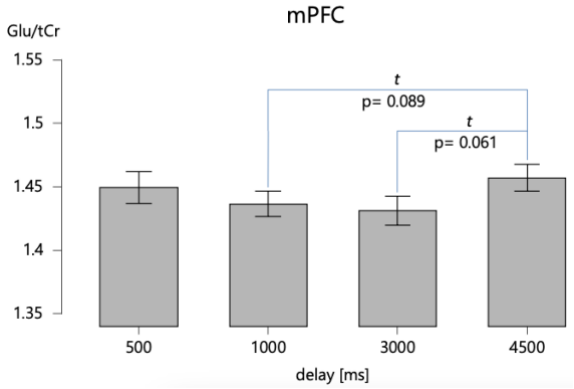

Figure 28. Glutamate concentration differences at different time delays after stimulation on 3 Tesla scanner.

No statistically significant effects of stimulation type was found ($F(1, 36) = 1.062$, $p = 0.310$, $\eta^2_p = 0.032$, $BF_{incl} = 0.095$) was found. Furthermore no statistically significant interactions between stimulation type and group ($F(1, 36) = 0.110$, $p = 0.742$, $\eta^2_p = 0.003$, $BF_{incl} = 0.247$), between stimulation type and sex ($F(1, 36) = 1.603$, $p = 0.215$, $\eta^2_p = 0.048$, $BF_{incl} = 0.276$), between delay and group ($F(3, 34) = 0.541$, $p = 0.655$, $\eta^2_p = 0.017$,

BF_{incl} = 0.067), between delay and sex (F(3, 34) = 0.426, p = 0.735, η^2_p = 0.013, BF_{incl} =0.083), between delay, group and sex (F(3, 34) = 0.673, p = 0.571, η^2_p = 0.021, BF_{incl} = 0.199), between stimulation type, delay and group (F(3, 34) = 0.638, p = 0.592, η^2_p = 0.020, BF_{incl} = 0.136), between stimulation type, delay and sex (F(3, 34) = 0.304, p = 0.822, η^2_p = 0.009, BF_{incl} = 0.166), between stimulation type, delay, group and sex (F(3, 34) = 0.615, p = 0.607, η^2_p = 0.019, BF_{incl} = 0.156) and between group and sex (F(1, 36) = 0.260, p = 0.614, η^2_p = 0.008, BF_{incl} = 0.390) were observed. Moreover, no statistically significant effects of group (F(1, 36) = 0.207, p = 0.652, η^2_p = 0.006, BF_{incl} = 0.220) and sex (F(1, 36) = 2.052, p = 0.162, η^2_p = 0.060, BF_{incl} = 1.828) were found.

After BOLD correction

After BOLD correction, the interaction between stimulation type, group, and sex (F(1, 35) = 4.679, p = 0.038, η^2_p = 0.131, BF_{incl} = 1.242) remained statistically significant. Contrasts showed that this was the result of a trend-level glutamate concentration difference between females and males (higher glutamate) in the dyslexic group after stimulation with false font strings (t(31) = -1.987, p = 0.056, d = -0.658) (see Figure 29). Moreover, the interaction between stimulation type and delay also remained statistically significant (F(3, 33) = 3.234, p = 0.026, η^2_p = 0.094, BF_{incl} = 0.189), induced by a trend-level lower glutamate concentration 1000 ms after words stimulation compared to 3000 ms (t(31) = -1.799, p = 0.082, d = -0.308) and 4500 ms (t(31) = -1.937, p = 0.062, d = -0.348) (see Figure 30). A trend-level interaction between stimulation type, delay, and age (F(3, 33) = 2.542, p = 0.061, η^2_p = 0.076), as well as between stimulation type, delay, and gray matter volume (F(3, 33) = 2.231, p = 0.090, η^2_p = 0.067), was observed. A statistically significant effect of participants' age was observed (F(1, 35) = 4.332, p = 0.046, η^2_p = 0.123, BF_{incl} = 2.551).

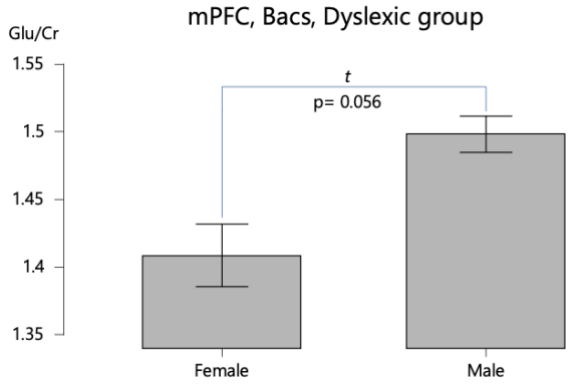

Figure 29. Difference in glutamate concentration between the control and dyslexic groups in the medial prefrontal cortex observed on a 3 Tesla scanner for data grouped by delay.

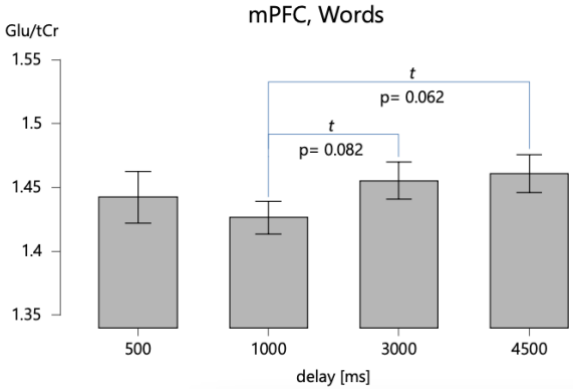

Figure 30. Difference in glutamate concentration at different time delays after word stimulation in the medial prefrontal cortex observed on a 3 Tesla scanner for data grouped by delay.

No statistically significant effects of delay ($F(3, 33) = 0.902$, $p = 0.443$, $\eta^2_p = 0.028$, $BF_{incl} = 0.029$) and stimulation type ($F(1, 35) = 0.175$, $p = 0.679$, $\eta^2_p = 0.006$, $BF_{incl} = 0.167$) were revealed. Furthermore, no statistically significant interactions between stimulation type and group ($F(1, 35) = 0.002$, $p = 0.967$, $\eta^2_p = 5.542 \cdot 10^{-5}$, $BF_{incl} = 0.215$), between stimulation type and sex ($F(1, 35) = 0.222$, $p = 0.641$, $\eta^2_p = 0.007$, $BF_{incl} = 0.505$), between delay and group ($F(3, 33) = 0.905$, $p = 0.442$, $\eta^2_p = 0.028$, $BF_{incl} = 0.122$), between delay and sex ($F(3, 33) = 0.396$, $p = 0.756$, $\eta^2_p = 0.013$, $BF_{incl} = 0.083$), between delay, group and sex ($F(3, 33) = 1.336$, $p = 0.267$, $\eta^2_p = 0.041$, $BF_{incl} = 0.308$), between stimulation type, delay and group ($F(3, 33) = 0.470$, $p = 0.704$, $\eta^2_p = 0.015$, $BF_{incl} = 0.060$), between stimulation type, delay and sex ($F(3, 33) = 0.422$, $p = 0.738$, $\eta^2_p = 0.013$, $BF_{incl} = 0.202$), between stimulation type, delay, group and sex ($F(3, 33) = 0.892$, $p = 0.449$, $\eta^2_p = 0.028$, $BF_{incl} = 0.258$) and between group and sex ($F(1, 35) = 0.182$, $p = 0.673$, $\eta^2_p = 0.006$, $BF_{incl} = 0.371$) were observed. Additionally, no statistically significant effects of group

(F(1, 35) = 0.245, p = 0.624, η^2_p = 0.008, BF_{incl} = 0.352) and sex (F(1, 35) = 2.264, p = 0.143, η^2_p = 0.068, BF_{incl} = 0.873) were found.

Left superior temporal sulcus

Analysis conducted for the left superior temporal sulcus showed a trend-level statistically significant three-way interaction between stimulation type, delay, and sex (F(3, 31) = 2.176, p = 0.097, η^2_p = 0.070, BF_{incl} = 1.051), with contrasts suggesting a trend-level difference between words and false font strings 500 ms after stimulation in females (t(29) = 2.015, p = 0.053, d = 0.317), and between words and false font strings 1000 ms after stimulation in males (t(29) = -1.802, p = 0.082, d = -0.315). A trend-level difference in males after false font string stimulation was also observed between 1000 ms and 3000 ms (t(29) = 1.785, p = 0.085, d = 0.334) (see Figure 31). Gray matter volume inside the analyzed voxel was statistically significant (F(1, 33) = 8.541, p = 0.007, η^2_p = 0.228, BF_{incl} = 6.243).

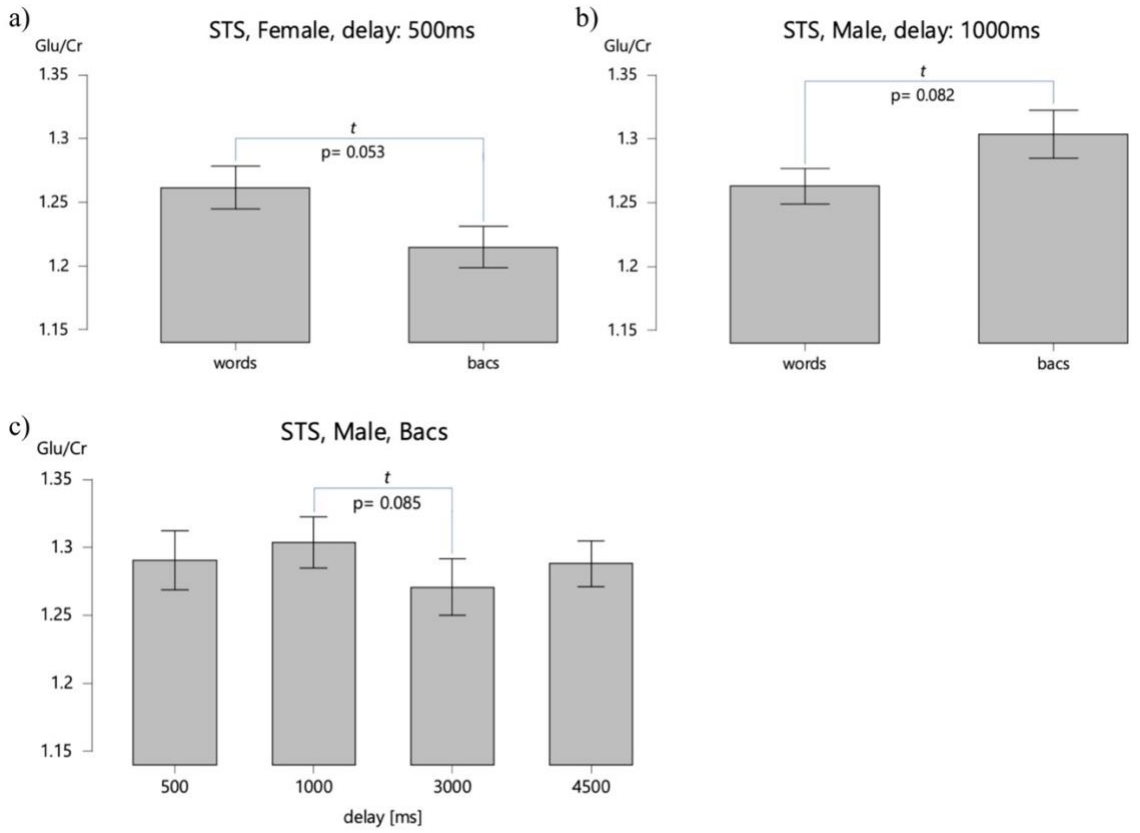

Figure 31. Trend-level interaction between stimulation type, delay, and sex: a) glutamate concentration differences in females after 500 ms of stimulation; b) glutamate concentration differences in males after 1000

ms of stimulation; c) glutamate concentration differences between delays in males after false font string stimulation on 3 Tesla scanner.

No statistically significant effects of delay ($F(3, 31) = 1.047$, $p = 0.376$, $\eta^2_p = 0.035$, $BF_{incl} = 0.026$) and stimulation type ($F(1, 33) = 1.249$, $p = 0.273$, $\eta^2_p = 0.041$, $BF_{incl} = 0.196$) were observed. No statistically significant interactions between stimulation type and group ($F(1, 33) = 0.098$, $p = 0.757$, $\eta^2_p = 0.003$, $BF_{incl} = 0.155$), between stimulation type and sex ($F(1, 33) = 1.456$, $p = 0.237$, $\eta^2_p = 0.048$, $BF_{incl} = 0.292$), between stimulation type, group and sex ($F(1, 33) = 2.313 \cdot 10^{-4}$, $p = 0.988$, $\eta^2_p = 7.976 \cdot 10^{-6}$, $BF_{incl} = 0.358$), between delay and group ($F(3, 31) = 0.743$, $p = 0.530$, $\eta^2_p = 0.025$, $BF_{incl} = 0.089$), between delay and sex ($F(3, 31) = 0.474$, $p = 0.701$, $\eta^2_p = 0.016$, $BF_{incl} = 0.072$), between delay, group and sex ($F(3, 31) = 0.535$, $p = 0.660$, $\eta^2_p = 0.018$, $BF_{incl} = 0.063$), between stimulation type and delay ($F(3, 31) = 0.470$, $p = 0.704$, $\eta^2_p = 0.016$, $BF_{incl} = 0.060$), between stimulation type, delay and group ($F(3, 31) = 0.218$, $p = 0.884$, $\eta^2_p = 0.007$, $BF_{incl} = 0.018$), between stimulation type, delay, group and sex ($F(3, 31) = 0.707$, $p = 0.550$, $\eta^2_p = 0.024$, $BF_{incl} = 0.001$) and between group and sex ($F(1, 33) = 0.524$, $p = 0.475$, $\eta^2_p = 0.018$, $BF_{incl} = 0.614$) were found. Moreover, no statistically significant effects of group ($F(1, 33) = 1.447$, $p = 0.239$, $\eta^2_p = 0.048$, $BF_{incl} = 0.662$) and sex ($F(1, 33) = 0.268$, $p = 0.609$, $\eta^2_p = 0.009$, $BF_{incl} = 0.554$) were revealed.

After BOLD correction

After apodization, a trend-level effect of group was observed ($F(1, 30) = 3.101$, $p = 0.090$, $\eta^2_p = 0.107$, $BF_{incl} = 0.598$), with higher glutamate concentration in the control group (see Figure 32). A statistically significant effect of gray matter volume was revealed ($F(1, 30) = 5.843$, $p = 0.023$, $\eta^2_p = 0.184$, $BF_{incl} = 2.843$).

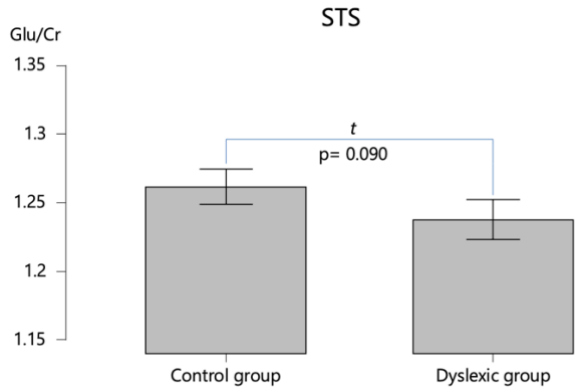

Figure 32. Difference in glutamate concentration between the control and dyslexic groups after BOLD correction in the left superior temporal sulcus, based on data grouped by delay acquired on a 3 Tesla scanner.

No statistically significant effects of delay ($F(3, 28) = 0.873$, $p = 0.459$, $\eta^2_p = 0.032$, $BF_{incl} = 0.033$) and stimulation type ($F(1, 30) = 1.650$, $p = 0.210$, $\eta^2_p = 0.060$, $BF_{incl} = 0.257$) were revealed. No statistically significant interactions between stimulation type and group ($F(1, 30) = 0.236$, $p = 0.631$, $\eta^2_p = 0.009$, $BF_{incl} = 0.302$), between stimulation type and sex ($F(1, 30) = 2.083$, $p = 0.161$, $\eta^2_p = 0.074$, $BF_{incl} = 0.513$), between stimulation type, group and sex ($F(1, 30) = 1.263$, $p = 0.271$, $\eta^2_p = 0.046$, $BF_{incl} = 0.338$), between delay and group ($F(3, 28) = 0.374$, $p = 0.772$, $\eta^2_p = 0.014$, $BF_{incl} = 0.076$), between delay and sex ($F(3, 28) = 0.566$, $p = 0.639$, $\eta^2_p = 0.021$, $BF_{incl} = 0.076$), between delay, group and sex ($F(3, 28) = 0.534$, $p = 0.660$, $\eta^2_p = 0.020$, $BF_{incl} = 0.141$), between stimulation type and delay ($F(3, 28) = 0.534$, $p = 0.660$, $\eta^2_p = 0.020$, $BF_{incl} = 0.111$), between stimulation type, delay and group ($F(3, 28) = 0.235$, $p = 0.872$, $\eta^2_p = 0.009$, $BF_{incl} = 0.195$), between stimulation type, delay and sex ($F(3, 28) = 2.065$, $p = 0.112$, $\eta^2_p = 0.074$, $BF_{incl} = 0.613$), between stimulation type, delay, group and sex ($F(3, 28) = 0.113$, $p = 0.952$, $\eta^2_p = 0.004$, $BF_{incl} = 0.105$) and between group and sex ($F(1, 30) = 1.768$, $p = 0.195$, $\eta^2_p = 0.064$, $BF_{incl} = 0.820$) were observed. No statistically significant effect of sex ($F(1, 30) = 1.320$, $p = 0.261$, $\eta^2_p = 0.048$, $BF_{incl} = 0.683$) was found.

3.2.3.5. Data analyzed using the dynamic-averaged approach

Medial prefrontal cortex

Analysis using the dynamic-averaged approach did not find a significant effect of delay ($F(3, 34) = 1.687$, $p = 0.175$, $\eta^2_p = 0.050$, $BF_{incl} = 0.099$) or stimulation type ($F(1, 36) = 0.666$, $p = 0.421$, $\eta^2_p = 0.020$, $BF_{incl} = 0.140$) in the medial prefrontal cortex.

A statistically significant interaction between stimulation type, group, and sex was revealed ($F(1, 36) = 5.253$, $p = 0.029$, $\eta^2_p = 0.141$, $BF_{incl} = 1.571$), driven by a trend-level difference in typical reading females between stimulation using words (lower glutamate concentration) and false font strings ($t(32) = -1.926$, $p = 0.063$, $d = -0.292$) (see Figure 33). Additionally, the interaction between stimulation type and delay was statistically significant ($F(3, 34) = 3.680$, $p = 0.015$, $\eta^2_p = 0.103$, $BF_{incl} = 0.177$), with lower glutamate levels 1000 ms after word stimulation compared to 4500 ms ($t(32) = -2.076$, $p = 0.046$, $d = -0.357$) and a trend-level higher glutamate concentration 500 ms after false font string stimulation compared to 3000 ms ($t(32) = 1.913$, $p = 0.065$, $d = 0.382$) (see Figure 34). Moreover, statistically significant interactions between stimulation type, delay, and age ($F(3, 34) = 3.198$, $p = 0.027$, $\eta^2_p = 0.091$), and between stimulation type, delay, and gray matter volume ($F(3, 34) = 3.258$, $p = 0.025$, $\eta^2_p = 0.092$), were observed. A significant effect of age was discovered ($F(1, 36) = 4.640$, $p = 0.039$, $\eta^2_p = 0.127$, $BF_{incl} = 3.119$).

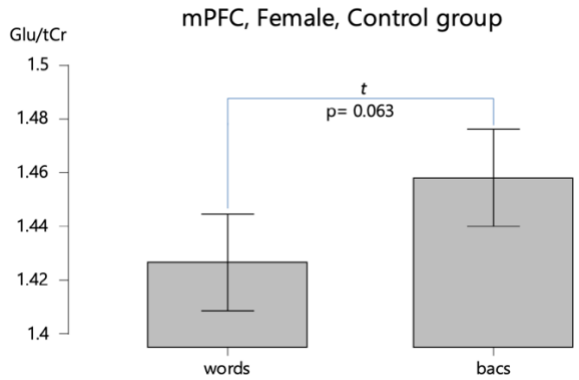

Figure 33. Difference in glutamate concentration in females from the control group between stimulation with words and false font strings (written in bacs font), based on data analyzed using the dynamic-averaged approach on 3 Tesla scanner.

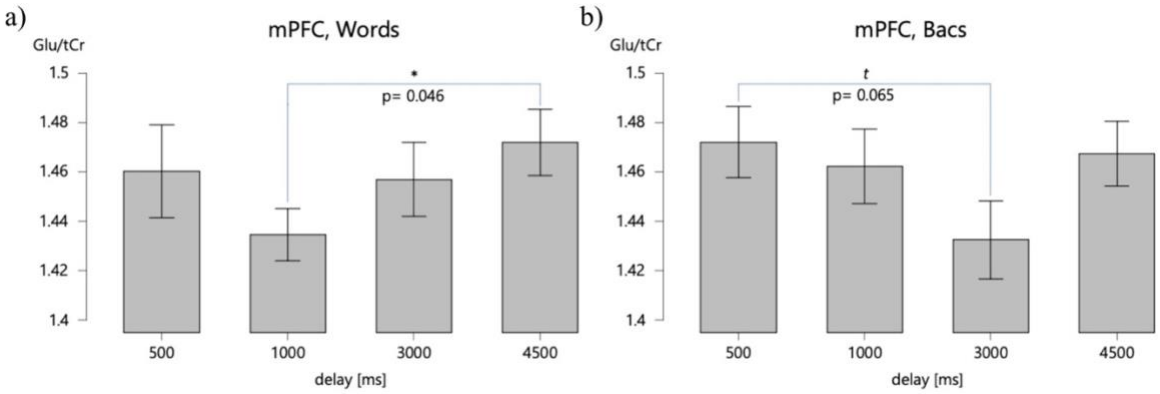

Figure 34. Difference in glutamate concentration on 3 Tesla scanner at different time delays: a) after word stimulation; b) after false font string stimulation, based on data analyzed using the dynamic-averaged approach in the medial prefrontal cortex.

Furthermore, no statistically significant interactions between stimulation type and group ($F(1, 36) = 6.475 \cdot 10^{-7}$, $p = 0.999$, $\eta^2_p = 2.023 \cdot 10^{-8}$, $BF_{incl} = 0.726$), between stimulation type and sex ($F(1, 36) = 1.780$, $p = 0.192$, $\eta^2_p = 0.053$, $BF_{incl} = 0.299$), between delay and group ($F(3, 34) = 0.453$, $p = 0.716$, $\eta^2_p = 0.014$, $BF_{incl} = 0.065$), between delay and sex ($F(3, 34) = 0.299$, $p = 0.826$, $\eta^2_p = 0.009$, $BF_{incl} = 0.061$), between delay, group and sex ($F(3, 34) = 0.953$, $p = 0.418$, $\eta^2_p = 0.029$, $BF_{incl} = 0.180$), between stimulation type, delay and group ($F(3, 34) = 0.615$, $p = 0.607$, $\eta^2_p = 0.019$, $BF_{incl} = 0.178$), between stimulation type, delay and sex ($F(3, 34) = 0.077$, $p = 0.972$, $\eta^2_p = 0.002$, $BF_{incl} = 0.073$), between stimulation type, delay, group and sex ($F(3, 34) = 0.341$, $p = 0.795$, $\eta^2_p = 0.011$, $BF_{incl} = 0.115$) and between group and sex ($F(1, 36) = 0.348$, $p = 0.559$, $\eta^2_p = 0.011$, $BF_{incl} = 0.505$) were observed. Additionally, no statistically significant effects of group ($F(1, 36) = 0.060$, $p = 0.809$, $\eta^2_p = 0.002$, $BF_{incl} = 0.263$) and sex ($F(1, 36) = 2.130$, $p = 0.154$, $\eta^2_p = 0.062$, $BF_{incl} = 0.915$) were found.

After BOLD correction

After the procedure aimed at correcting the potential influence of the BOLD effect on spectra, a statistically significant interaction between stimulation type, group, and sex remained significant ($F(1, 35) = 5.761$, $p = 0.023$, $\eta^2_p = 0.157$, $BF_{incl} = 1.011$), with a trend-level difference between females and males (stronger glutamate response) diagnosed with dyslexia after false font string stimulation ($t(31) = -1.727$, $p = 0.094$, $d = -0.607$). A trend-level difference was also observed in females from the control group after words (lower glutamate concentration) and false font string stimulation

(t(31) = -1.807, p = 0.080, d = -0.270) (see Figure 35). A statistically significant interaction was observed between stimulation type and delay ((F(3, 33) = 3.402, p = 0.021, η^2_p = 0.099, BF_{incl} = 0.117), caused by the difference in glutamate concentration after words stimulation between 1000 ms and 4500 ms (t(31) = -2.275, p = 0.030, d = -0.378), with higher glutamate after 4500 ms (see Figure 36). Additionally, a statistically significant interaction was found between stimulation type, delay, and age (F(3, 33) = 2.865, p = 0.041, η^2_p = 0.085), and a trend-level interaction between stimulation type, delay, and gray matter volume (F(3, 33) = 2.596, p = 0.057, η^2_p = 0.077). A statistically significant effect of participants' age was revealed (F(1, 35) = 5.130, p = 0.031, η^2_p = 0.142, BF_{incl} = 3.480).

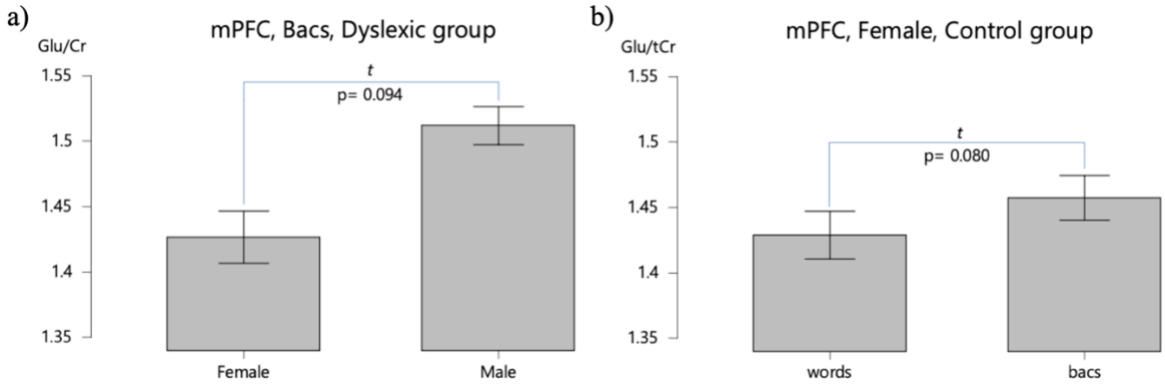

Figure 35. Glutamate differences on 3 Tesla scanner: a) between females and males in the dyslexic group after false font string stimulation; b) between stimulation with words and false font strings in typical reading females.

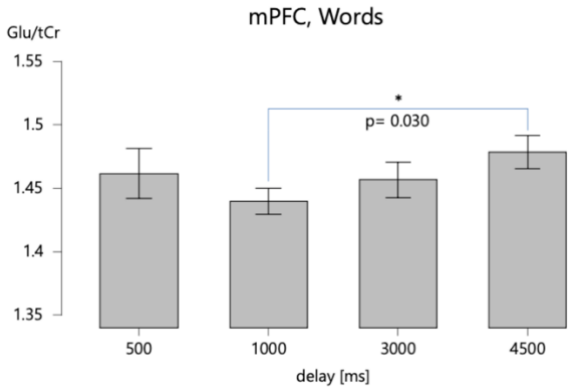

Figure 36. Difference in glutamate concentration at different time delays after word stimulation in the medial prefrontal cortex, after BOLD correction, observed on a 3 Tesla scanner for data analyzed using the dynamic-averaged approach.

The effects of delay (F(3, 33) = 1.424, p = 0.241, η^2_p = 0.044, BF_{incl} = 0.097) and stimulation type (F(1, 35) = 0.668, p = 0.420, η^2_p = 0.021, BF_{incl} = 0.169) were not significant.

No statistically significant interactions between stimulation type and group ($F(1, 35) = 0.020$, $p = 0.887$, $\eta^2_p = 6.566 \cdot 10^{-4}$, $BF_{incl} = 0.217$), between stimulation type and sex ($F(1, 35) = 0.108$, $p = 0.744$, $\eta^2_p = 0.003$, $BF_{incl} = 0.245$), between delay and group ($F(3, 33) = 0.569$, $p = 0.636$, $\eta^2_p = 0.018$, $BF_{incl} = 0.068$), between delay and sex ($F(3, 33) = 0.205$, $p = 0.893$, $\eta^2_p = 0.007$, $BF_{incl} = 0.039$), between delay, group and sex ($F(3, 33) = 1.177$, $p = 0.323$, $\eta^2_p = 0.037$, $BF_{incl} = 0.244$), between stimulation type, delay and group ($F(3, 33) = 0.524$, $p = 0.667$, $\eta^2_p = 0.017$, $BF_{incl} = 0.172$), between stimulation type, delay and sex ($F(3, 33) = 0.179$, $p = 0.910$, $\eta^2_p = 0.006$, $BF_{incl} = 0.129$), between stimulation type, delay, group and sex ($F(3, 33) = 0.547$, $p = 0.651$, $\eta^2_p = 0.017$, $BF_{incl} = 0.102$) and between group and sex ($F(1, 35) = 0.190$, $p = 0.666$, $\eta^2_p = 0.006$, $BF_{incl} = 0.322$) were observed. No statistically significant effects of group ($F(1, 35) = 0.183$, $p = 0.672$, $\eta^2_p = 0.006$, $BF_{incl} = 0.341$) and sex ($F(1, 35) = 2.253$, $p = 0.143$, $\eta^2_p = 0.068$, $BF_{incl} = 1.326$) were found.

Left superior temporal sulcus

In the left superior temporal cortex, the effects of delay ($F(3, 31) = 0.864$, $p = 0.463$, $\eta^2_p = 0.029$, $BF_{incl} = 0.100$) and stimulation type ($F(1, 33) = 1.759$, $p = 0.195$, $\eta^2_p = 0.057$, $BF_{incl} = 0.196$) were not significant.

A statistically significant interaction between stimulation type, delay, and sex was observed ($F(3, 31) = 3.794$, $p = 0.013$, $\eta^2_p = 0.116$, $BF_{incl} = 2.141$), with higher glutamate concentration after words compared to false font string stimulation in females at 500 ms ($t(29) = 2.083$, $p = 0.046$, $d = 0.313$). A significant difference between stimulation types was also found for males at 1000 ms ($t(29) = -2.450$, $p = 0.021$, $d = -0.331$), with higher glutamate levels after false font string stimulation compared to words. Additionally, for females stimulated with false font strings, a difference between 500 ms and 3000 ms ($t(29) = -2.074$, $p = 0.047$, $d = -0.431$), and 4500 ms ($t(29) = -2.043$, $p = 0.050$, $d = -0.439$) was detected, with lower glutamate after 500 ms. A trend-level difference was observed between 1000 ms (lower glutamate) and 3000 ms ($t(29) = -1.765$, $p = 0.088$, $d = -0.237$), and 4500 ms ($t(29) = -1.840$, $p = 0.076$, $d = -0.245$). Moreover, males stimulated with false font strings showed a stronger glutamate response after 1000 ms compared to 3000 ms ($t(29) = 2.547$, $p = 0.016$, $d = 0.333$) (see Figure 37). A trend-level interaction between stimulation type and gray matter volume was demonstrated ($F(1, 33) = 3.077$, $p = 0.090$,

$\eta^2_p = 0.096$). A statistically significant effect of gray matter volume within the voxel was reported ($F(1, 33) = 7.703$, $p = 0.010$, $\eta^2_p = 0.210$, $BF_{incl} = 5.415$).

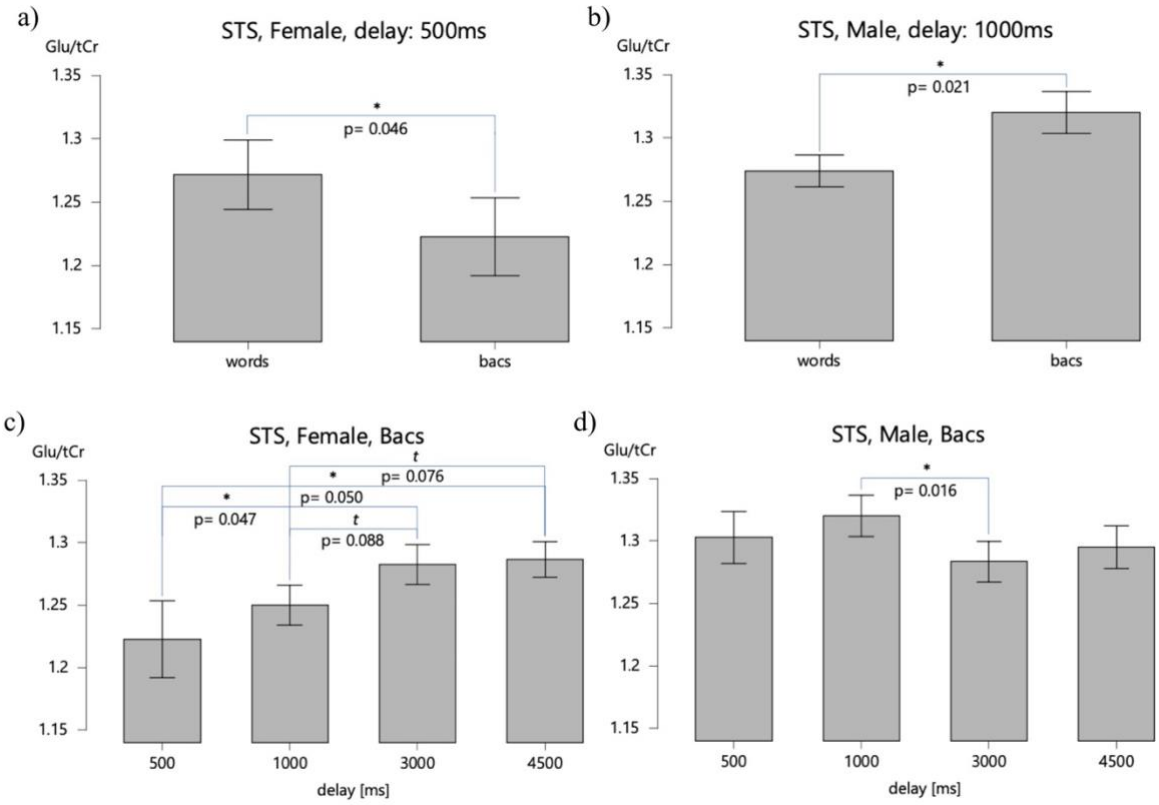

Figure 37. Difference in glutamate concentration in statistically significant interaction between stimulation type, delay, and sex between on 3 Tesla scanner: a) words and false font strings stimulation in females after 500 ms; b) words and false font strings stimulation in males after 1000 ms; c) different delays after false font strings stimulations in females; d) different delays after false font strings stimulations in males, in left superior temporal sulcus in dynamic-averaged approach.

Furthermore, no statistically significant interactions between stimulation type and group ($F(1, 33) = 0.401$, $p = 0.532$, $\eta^2_p = 0.014$, $BF_{incl} = 0.241$), between stimulation type and sex ($F(1, 33) = 1.764$, $p = 0.194$, $\eta^2_p = 0.057$, $BF_{incl} = 0.459$), between stimulation type, group and sex ($F(1, 33) = 2.952 \cdot 10^{-5}$, $p = 0.996$, $\eta^2_p = 1.018 \cdot 10^{-6}$, $BF_{incl} = 0.867$), between delay and group ($F(3, 31) = 0.008$, $p = 0.999$, $\eta^2_p = 2.764 \cdot 10^{-4}$, $BF_{incl} = 0.094$), between delay and sex ($F(3, 31) = 0.677$, $p = 0.569$, $\eta^2_p = 0.023$, $BF_{incl} = 0.088$), between delay, group and sex ($F(3, 31) = 0.800$, $p = 0.497$, $\eta^2_p = 0.027$, $BF_{incl} = 0.180$), between stimulation type and delay ($F(3, 31) = 0.835$, $p = 0.478$, $\eta^2_p = 0.028$, $BF_{incl} = 0.064$), between stimulation type, delay and group ($F(3, 31) = 0.326$, $p = 0.807$, $\eta^2_p = 0.011$, $BF_{incl} = 0.216$), between stimulation type, delay, group and sex ($F(3, 31) = 0.847$, $p = 0.472$, $\eta^2_p = 0.028$,

BF_{incl} = 0.286) and between group and sex (F(1, 33) = 0.017, p = 0.898, $\eta^2_p = 5.804 \cdot 10^{-4}$, BF_{incl} = 0.510) were revealed. No statistically significant effects of group (F(1, 33) = 0.240, p = 0.628, η^2_p = 0.008, BF_{incl} = 0.580) and sex (F(1, 33) = 0.010, p = 0.919, $\eta^2_p = 3.607 \cdot 10^{-4}$, BF_{incl} = 0.775) were found.

After BOLD Correction

After BOLD correction, the interaction between stimulation type, delay, and sex remained statistically significant (F(3, 28) = 3.413, p = 0.021, η^2_p = 0.116, BF_{incl} = 1.489), driven by the difference between words (higher glutamate concentration) and false font strings for females after 500 ms (t(26) = 2.807, p = 0.009, d = 0.438) and between words (lower glutamate concentration) and false font strings for males after 1000 ms (t(26) = -3.002, p = 0.006, d = -0.399). Moreover, females after false font string stimulation showed different glutamate concentrations between 500 ms, when glutamate levels were lower compared to 3000 ms (t(26) = -2.725, p = 0.011, d = -0.592), 4500 ms (t(26) = -2.383, p = 0.025, d = -0.581), and a trend-level difference to 1000 ms (t(26) = -1.770, p = 0.088, d = -0.369). Furthermore, males after false font string stimulation demonstrated differences between delays at 1000 ms (higher glutamate concentration) and 3000 ms (t(26) = 3.605, p = 0.001, d = 0.430), and a trend-level difference to 4500 ms (t(26) = 1.760, p = 0.090, d = 0.243). A trend-level difference at 500 ms after false font string stimulation was observed between females and males (t(26) = -1.870, p = 0.073, d = -0.618) (see Figure 38). A trend-level interaction between stimulation type, delay, and age was revealed (F(3, 28) = 2.345, p = 0.079, η^2_p = 0.083). In addition, a trend-level effect of gray matter volume was found (F(1, 30) = 4.159, p = 0.052, η^2_p = 0.138, BF_{incl} = 1.832).

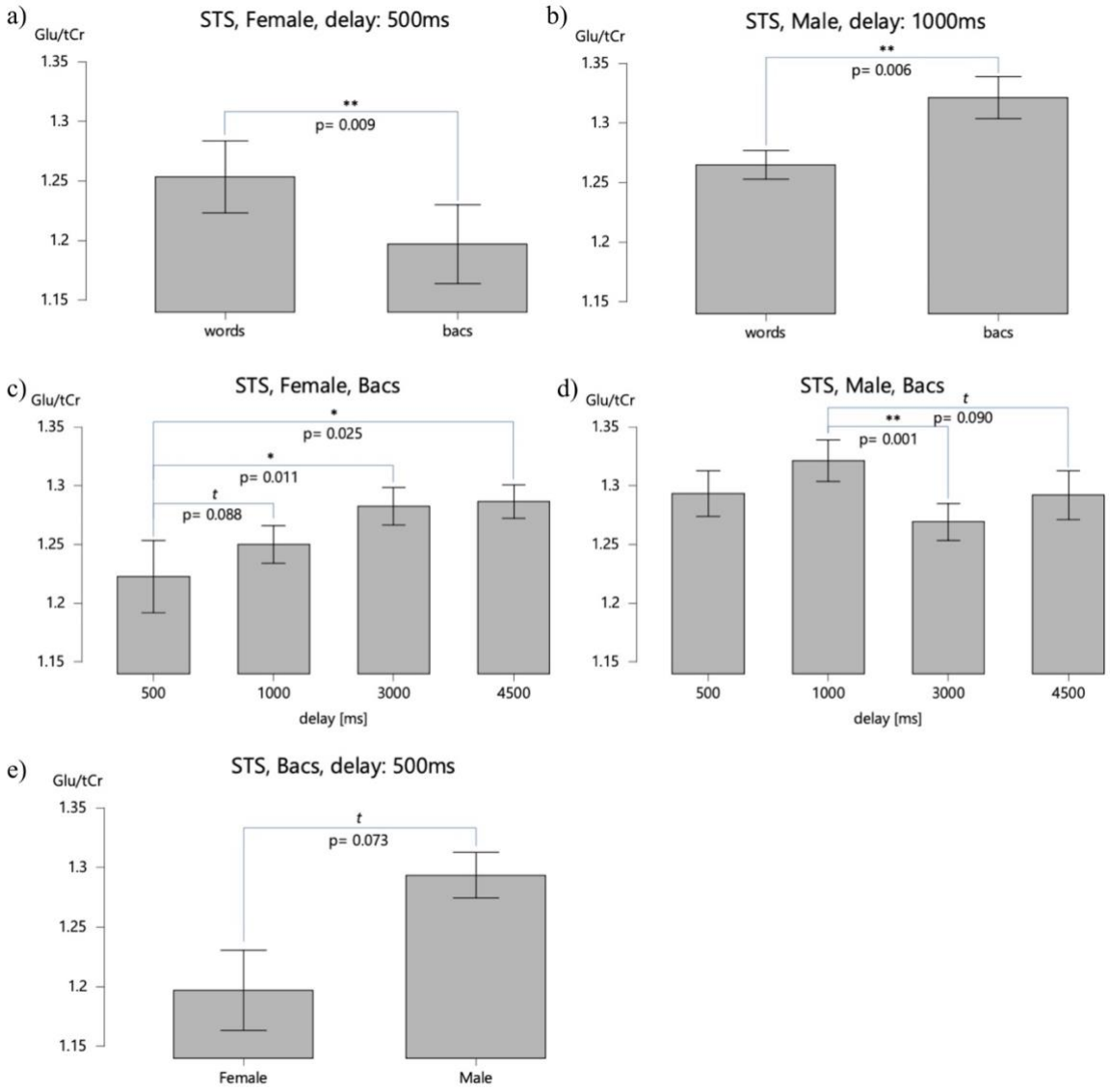

Figure 38. Glutamate concentration differences on 3 Tesla scanner: a) between words and false font string stimulation in females after 500 ms; b) between words and false font string stimulation in males after 1000 ms; c) differences between delays after false font string stimulation in females; d) differences between delays after false font string stimulation in males; e) between females and males at 500 ms after false font string stimulation, in the left superior temporal sulcus, using the dynamic-averaged approach after BOLD correction.

No statistically significant effect of delay ($F(3, 28) = 0.501$, $p = 0.683$, $\eta^2_p = 0.019$, $BF_{incl} = 0.118$) and stimulation type ($F(1, 30) = 1.337$, $p = 0.258$, $\eta^2_p = 0.049$, $BF_{incl} = 0.186$) were revealed.

No statistically significant interactions between stimulation type and group ($F(1, 30) = 0.076$, $p = 0.785$, $\eta^2_p = 0.003$, $BF_{incl} = 0.274$), between stimulation type and sex ($F(1, 30) = 2.155$, $p = 0.154$, $\eta^2_p = 0.077$, $BF_{incl} = 0.548$), between stimulation type, group

and sex ($F(1, 30) = 0.230$, $p = 0.635$, $\eta^2_p = 0.009$, $BF_{incl} = 0.378$), between delay and group ($F(3, 28) = 0.086$, $p = 0.968$, $\eta^2_p = 0.003$, $BF_{incl} = 0.100$), between delay and sex ($F(3, 28) = 1.069$, $p = 0.367$, $\eta^2_p = 0.039$, $BF_{incl} = 0.186$), between delay, group and sex ($F(3, 28) = 1.168$, $p = 0.327$, $\eta^2_p = 0.043$, $BF_{incl} = 0.752$), between stimulation type and delay ($F(3, 28) = 0.739$, $p = 0.532$, $\eta^2_p = 0.028$, $BF_{incl} = 0.187$), between stimulation type, delay and group ($F(3, 28) = 0.355$, $p = 0.786$, $\eta^2_p = 0.013$, $BF_{incl} = 0.125$), between stimulation type, delay, group and sex ($F(3, 28) = 1.212$, $p = 0.311$, $\eta^2_p = 0.045$, $BF_{incl} = 1.655$) and between group and sex ($F(1, 30) = 0.148$, $p = 0.704$, $\eta^2_p = 0.006$, $BF_{incl} = 0.642$) were observed. No statistically significant effects of group ($F(1, 30) = 0.593$, $p = 0.448$, $\eta^2_p = 0.022$, $BF_{incl} = 0.574$) and sex ($F(1, 30) = 0.146$, $p = 0.705$, $\eta^2_p = 0.006$, $BF_{incl} = 0.615$) were found.

3.3. Experiment 3 – comparing data quality between the 7T and the 3T scanners

The quality of data from the 7T and 3T scanners was compared based on parameters automatically calculated by the FSL-MRS software. Data were averaged across all spectra, with 320 signals per dataset from each voxel. The brain regions, the medial prefrontal cortex and the left superior temporal sulcus, were analyzed separately. The comparison was based on the signal-to-noise ratio (SNR), percentage Cramer-Rao lower bound (%CRLB), full width at half maximum (FWHM) for two metabolites: glutamate and N-acetylaspartic acid (NAA), and linewidth across the entire spectrum. Data were compared for the same participants who took part in both functional magnetic resonance spectroscopy sessions on both scanners. As a result, data from the medial prefrontal cortex were analyzed for 34 participants: 18 from the control group (11 females, 7 males), and 16 from the dyslexic group (8 females, 8 males). For the left superior temporal sulcus, data from 30 participants were analyzed: 17 from the control group (10 females, 7 males), and 13 from the dyslexic group (6 females, 7 males).

Linewidth and FWHM, expressed in Hertz (Hz), differed between the 7T and 3T scanners due to the varying Larmor frequencies of the two scanners. As a result, these values required recalculation to ensure consistency across the scanners. The results for the 3T scanner were adjusted by a factor corresponding to the difference in Larmor frequency between the two scanners. All presented results were corrected for the Larmor frequency difference.

3.3.1. Linewidth

Linewidth is a parameter that describes the width of the peaks in the spectrum. Narrower peaks are easier to distinguish, which implies that a lower linewidth value indicates better spectral quality. In other words, a smaller linewidth typically suggests higher resolution and more accurate identification of metabolite peaks, leading to more reliable data analysis. Linewidth may be influenced by shimming quality, which ensures the homogeneity of the magnetic field. Poor shimming can lead to broader peaks, resulting in higher linewidth values and potentially reducing spectral resolution.

In the medial prefrontal cortex, a statistically significant difference in linewidth between the 7T and 3T scanners was revealed ($t(33) = 3.589$, $p = 0.001$, $d = 0.615$), with a higher

mean linewidth on the 7T scanner 11.082 (SD = 2.982, Min = 6.120, Max = 17.850) compared to the 3T scanner 8.966 (SD = 1.885, Min = 5.540, Max = 14.210) (see Figure 39).

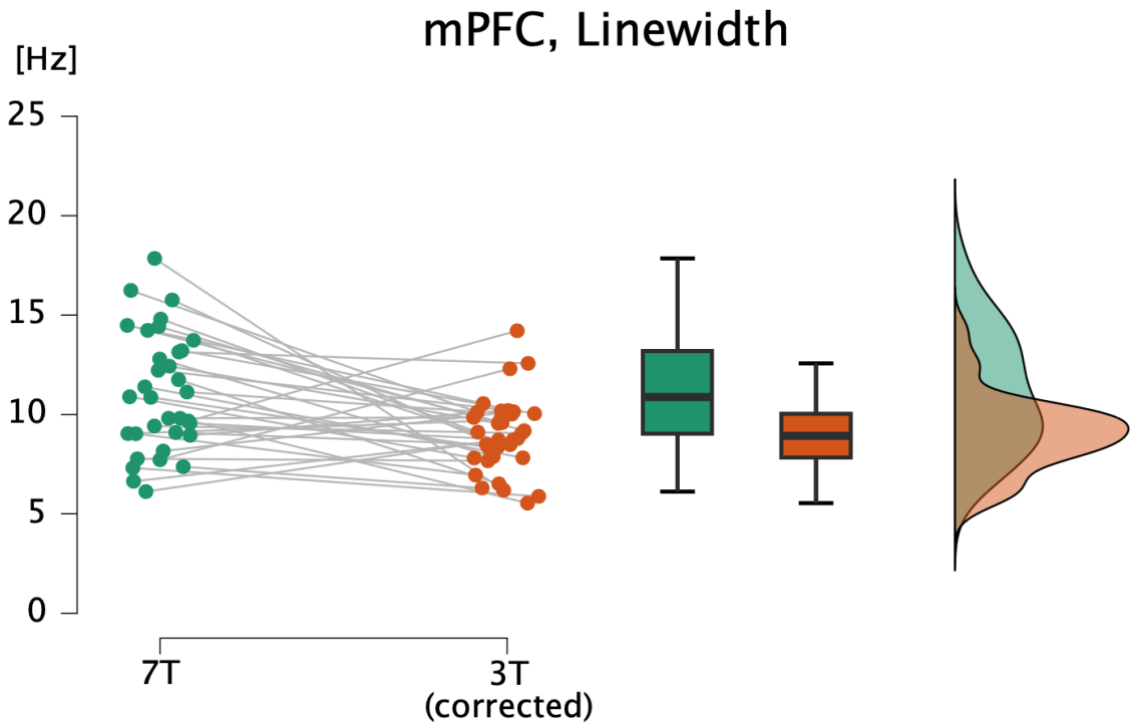

Figure 39. Linewidth comparison for data from the medial prefrontal cortex between the 7T and 3T scanners. Statistical comparison showed a significantly higher mean linewidth on the 7T scanner.

Linewidth in the left superior temporal sulcus did not statistically differ between scanners ($t(29) = 1.495$, $p = 0.146$, $d = 0.273$), with the mean linewidth for the 7T scanner at 10.760 (SD = 3.090, Min = 6.730, Max = 18.740) and for the 3T scanner (after correction) at 10.110 (SD = 2.102, Min = 7.230, Max = 16.180) (see Figure 40).

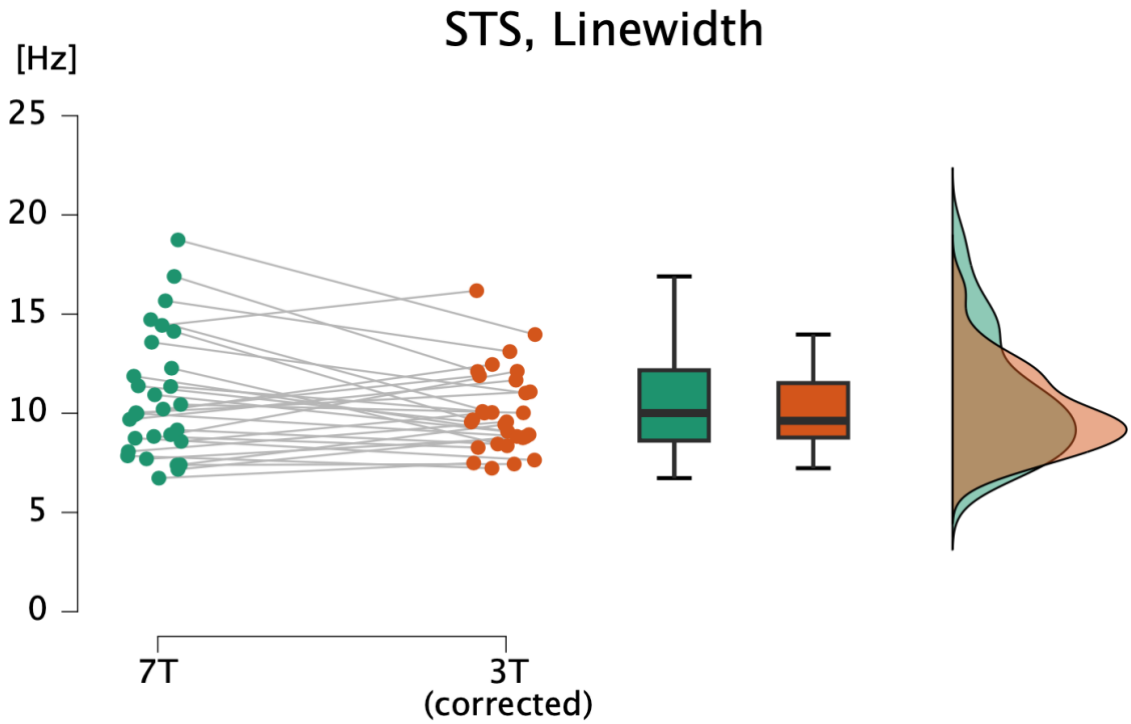

Figure 40. Linewidth comparison for data from the left superior temporal sulcus between the 7T and 3T scanners. No statistically significant difference was observed between the two scanners.

3.3.2. Signal-to-noise ratio (SNR)

The Signal-to-noise ratio (SNR) is a parameter calculated as the ratio of the amplitude of the main metabolite peak to the standard deviation of the noise, which is measured in a spectral area where no metabolite signals are present. It is expressed as:

$$SNR = \frac{\mu_{signal}}{\sigma_{noise}}$$

Where μ_{signal} is the mean amplitude of the main metabolite peak, and σ_{noise} is the standard deviation of the noise in a region of the spectrum where no metabolite signals are observed. A higher SNR indicates better data quality, as metabolite signals are easier to distinguish from noise.

Glutamate

In the control brain region, the SNR for glutamate was significantly different between the 7T and 3T scanners, with a higher SNR for the high-field scanner ($t(33) = 6.310$, $p < 0.001$, $d = 1.082$). The mean SNR for the 7T scanner was 27.109 (SD = 5.834, Min = 15.690,

Max = 42.780), while for the 3T scanner it was 20.425 (SD = 5.047, Min = 10.750, Max = 33.030) (see Figure 41).

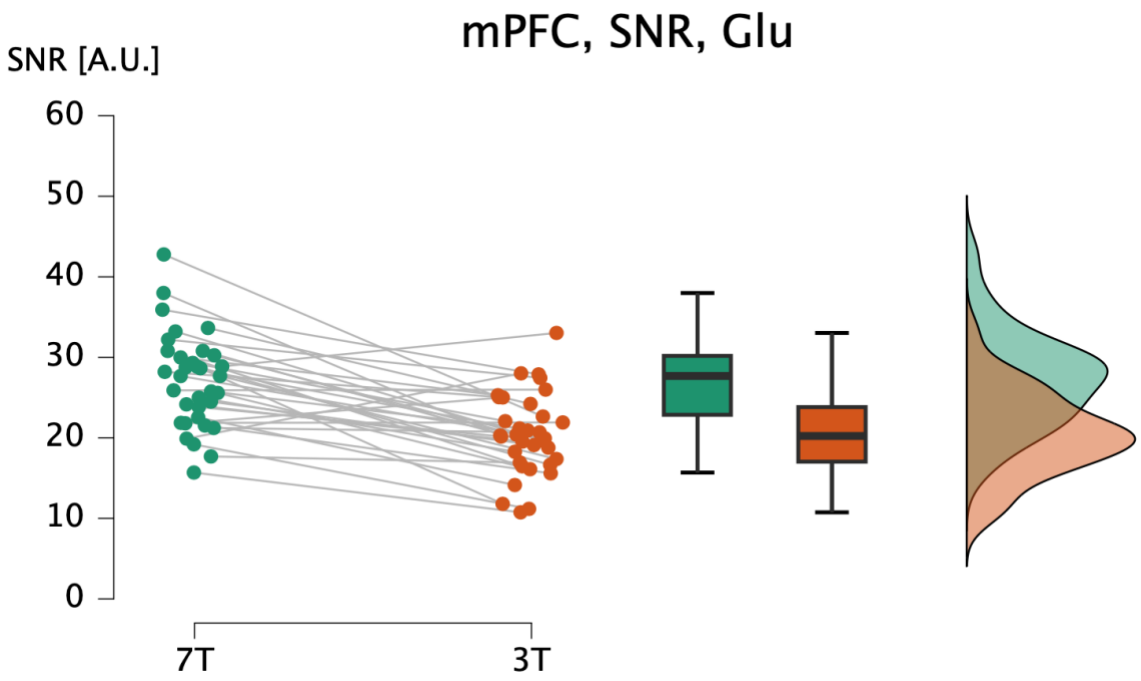

Figure 41. Signal-to-noise ratio (SNR) comparison for glutamate data from the medial prefrontal cortex between the 7T and 3T scanners. Statistical analysis showed a significantly higher mean SNR for glutamate on the 7T scanner.

No statistically significant difference in glutamate SNR was observed in the left superior temporal sulcus ($t(29) = 0.692$, $p = 0.494$, $d = 0.126$). The mean SNR for the 7T scanner was 20.972 (SD = 7.182, Min = 7.820, Max = 36.600), and for the 3T scanner it was 20.135 (SD = 5.307, Min = 11.140, Max = 30.540) (see Figure 42).

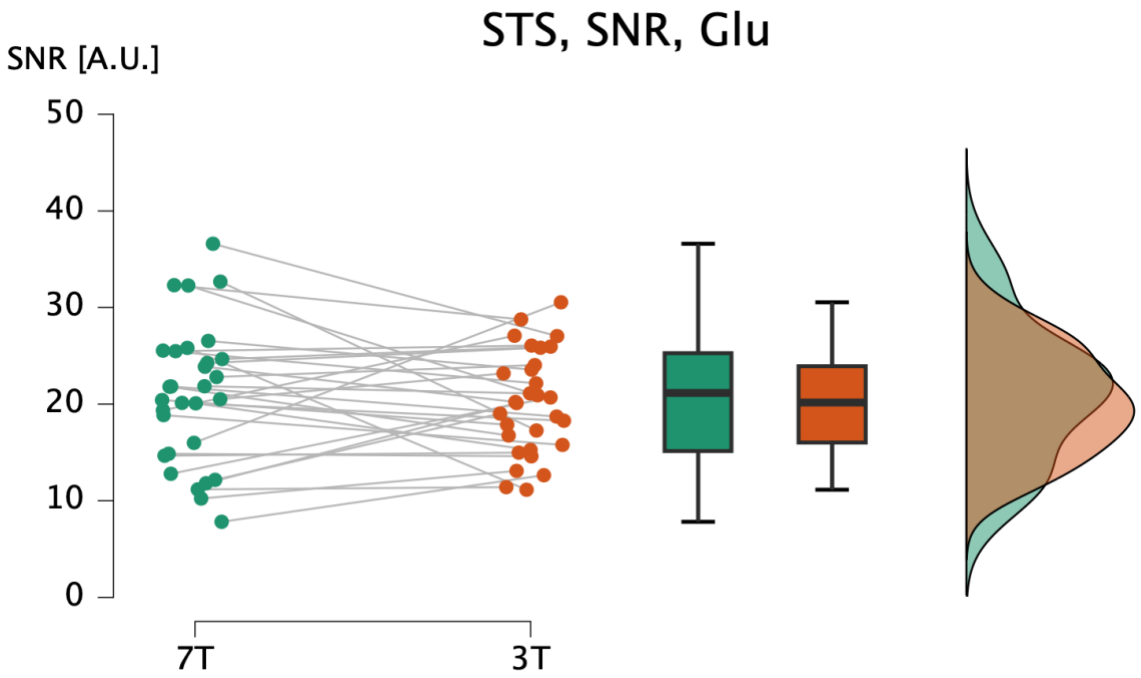

Figure 42. Signal-to-noise ratio (SNR) comparison for glutamate data from the left superior temporal sulcus between the 7T and 3T scanners. No statistically significant difference was observed between the two scanners.

NAA

No statistically significant difference in SNR for NAA was found in the medial prefrontal cortex between the two scanners ($t(33) = 1.602$, $p = 0.119$, $d = 0.275$). The mean SNR for the 7T scanner was 71.392 (SD = 22.309, Min = 36.900, Max = 134.330), and for the 3T scanner it was 64.675 (SD = 18.384, Min = 31.190, Max = 105.990) (see Figure 43).

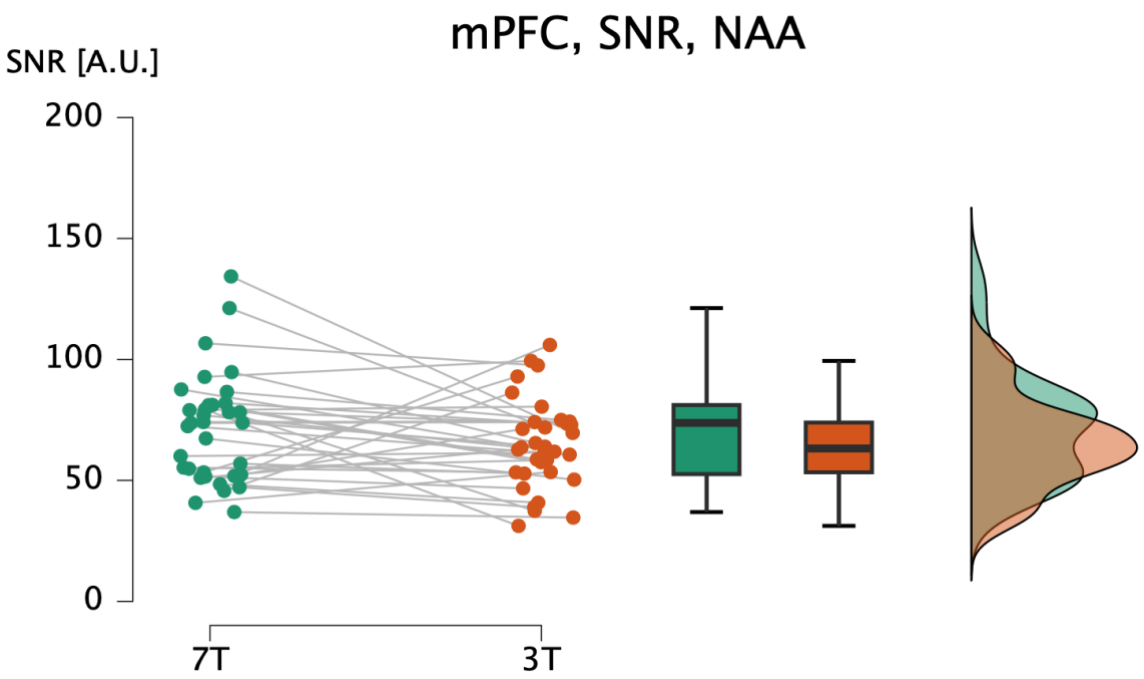

Figure 43. Signal-to-noise ratio (SNR) comparison for NAA data from the medial prefrontal cortex between the 7T and 3T scanners. No statistically significant difference was observed between the two scanners.

In the left superior temporal sulcus, no statistically significant differences between scanners were observed (t(29) = -0.071, p = 0.944, d = -0.013). For the high-field scanner, the mean SNR for NAA was 78.312 (SD = 27.375, Min = 22.160, Max = 143.830), and for the 3T scanner, it was 78.576 (SD = 16.052, Min = 44.970, Max = 113.190) (see Figure 44).

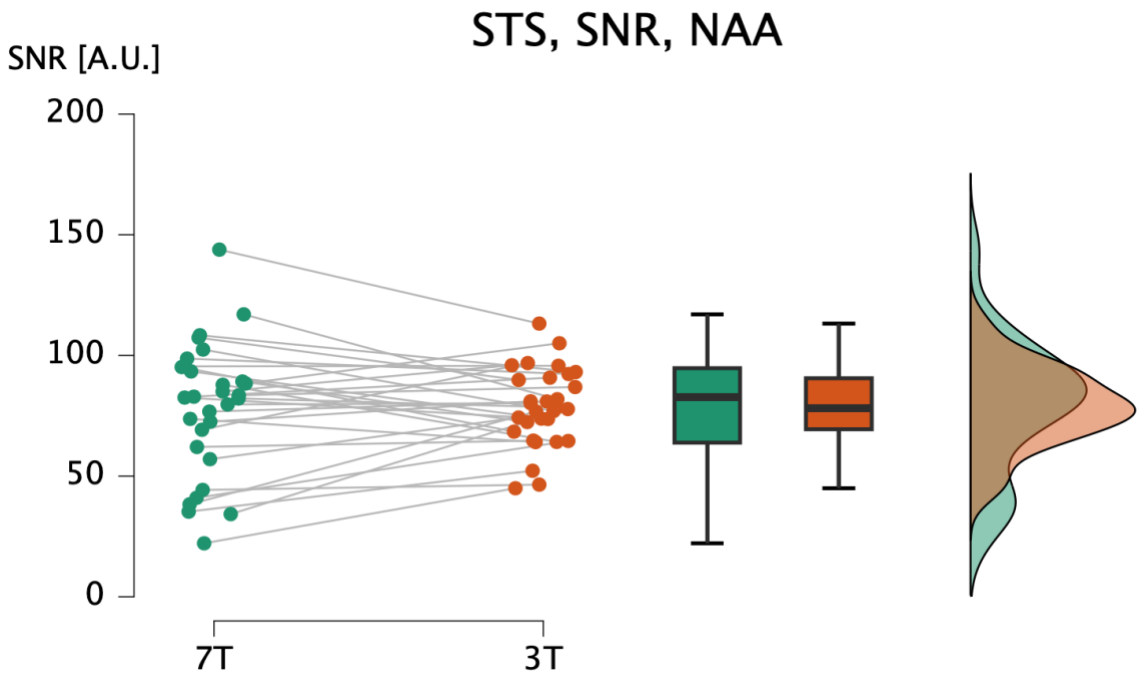

Figure 44. Signal-to-noise ratio (SNR) comparison for NAA data from the left superior temporal sulcus between the 7T and 3T scanners. No statistically significant difference was observed between the two scanners.

3.3.3. Percentage Cramer-Rao lower bound (%CRLB)

The Percentage Cramer-Rao Lower Bound (%CRLB) is a measure of the precision of the estimated metabolite concentration. A lower %CRLB indicates a more reliable and accurate estimate, reflecting lower estimation error. A higher %CRLB suggests a less accurate estimate, often due to weaker signals or increased noise in the data.

Glutamate

A statistically significant difference in %CRLB for glutamate between the 7T and 3T scanners in the medial prefrontal cortex was found. Due to a significant deviation from normality (Shapiro-Wilk test: $p = 0.003$), a nonparametric Wilcoxon signed-rank test was used ($Z = -4.941$, $p < 0.001$, effect size $r = -0.971$), with a lower %CRLB for the high-field scanner: the mean for the 7T scanner was 2.064 (SD = 0.282, Min = 1.630, Max = 2.730), and for the 3T scanner it was 2.721 (SD = 0.385, Min = 2.250, Max = 3.920) (see Figure 45).

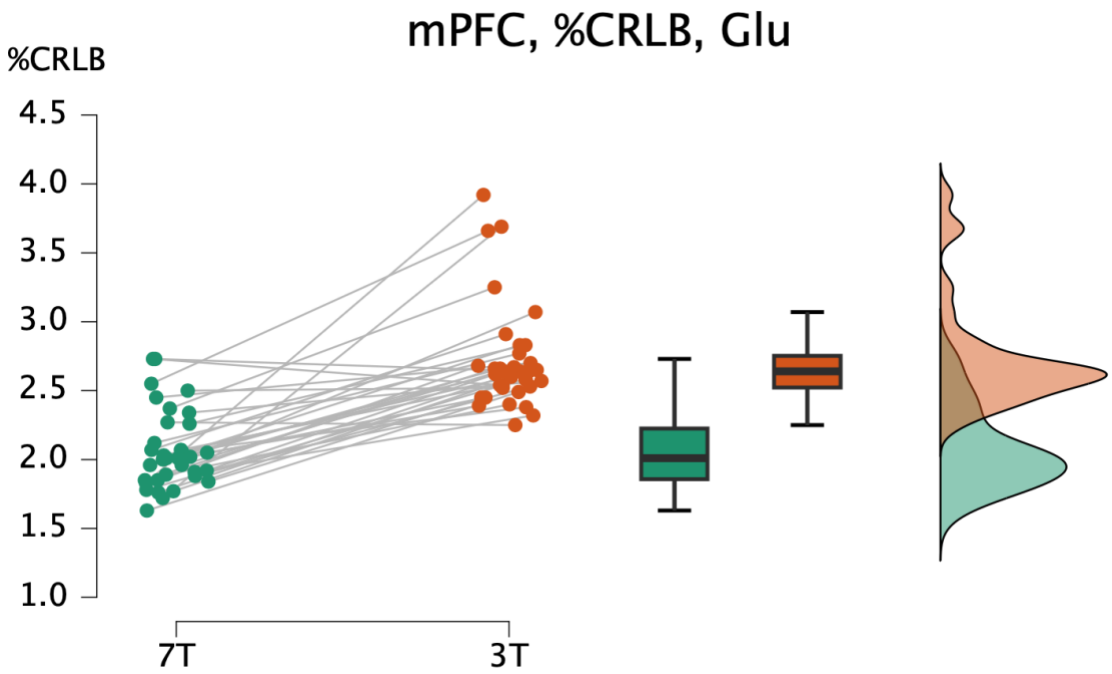

Figure 45. Percentage Cramer-Rao Lower Bound %CRLB comparison for glutamate data from the medial prefrontal cortex between the 7T and 3T scanners. Statistical analysis showed a significantly lower mean %CRLB for glutamate on the 7T scanner.

In the left superior temporal sulcus, a statistically significant difference between scanners was observed (Wilcoxon signed-rank test: $Z = -2.499$, $p = 0.011$, $r = -0.523$; Shapiro-Wilk test: $p = 0.033$), with better quality for the 7T scanner. The mean %CRLB for the 7T scanner was 3.075 (SD = 0.900, Min = 2.030, Max = 5.920), and for the 3T scanner it was 3.742 (SD = 1.177, Min = 2.390, Max = 7.430) (see Figure 46).

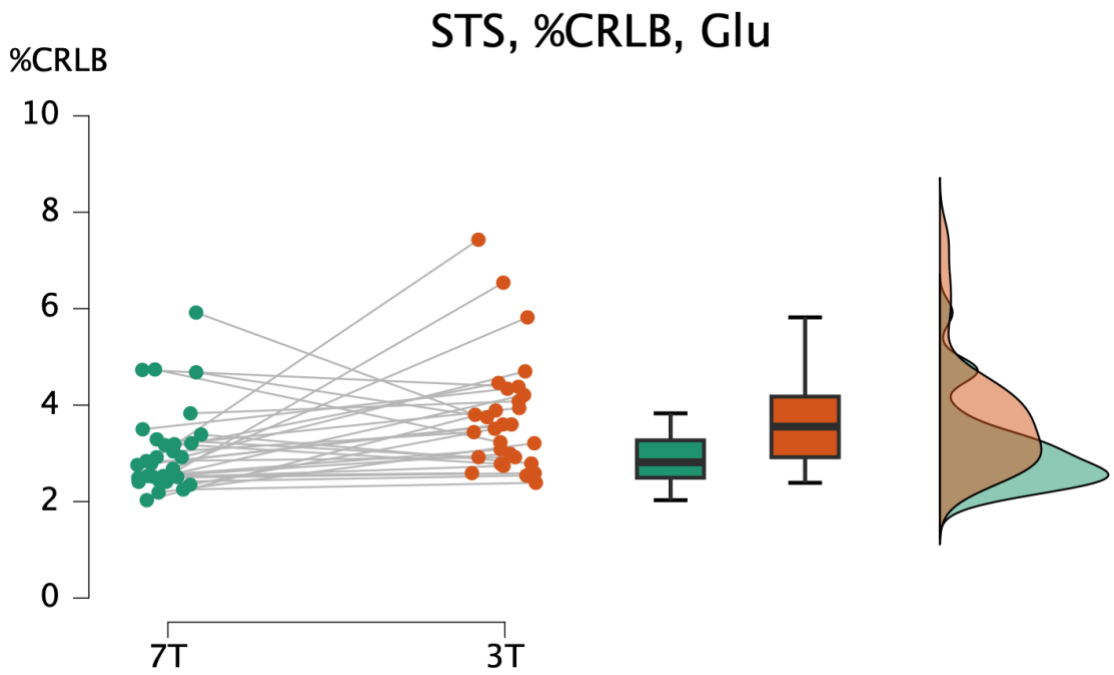

Figure 46. Percentage Cramer-Rao Lower Bound %CRLB comparison for glutamate data from the left superior temporal sulcus between the 7T and 3T scanners. Statistical analysis showed a significantly lower mean %CRLB for glutamate on the 7T scanner.

NAA

A statistically significant difference in %CRLB for NAA was found between the analyzed scanners ($t(33) = -2.757$, $p = 0.009$, $d = -0.473$), with a higher result for the 3T scanner. The mean %CRLB for the 7T scanner was 1.639 (SD = 0.388, Min = 1.080, Max = 2.630), and for the 3T scanner it was 1.907 (SD = 0.423, Min = 1.190, Max = 2.940) (see Figure 47).

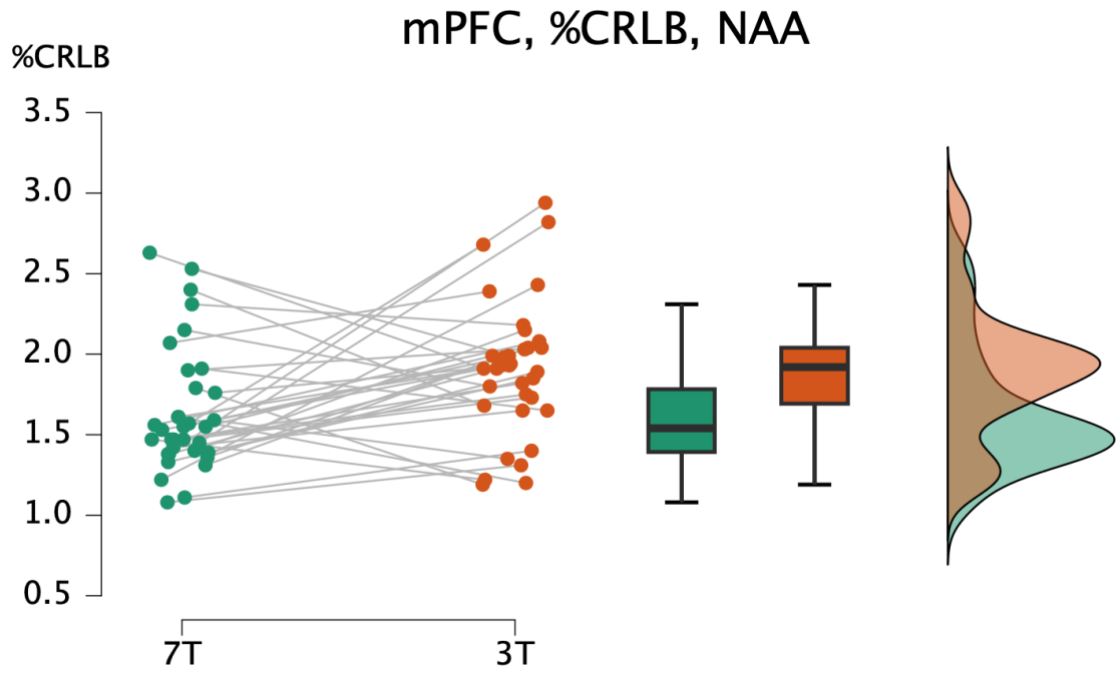

Figure 47. Percentage Cramer-Rao Lower Bound %CRLB comparison for NAA data from the medial prefrontal cortex between the 7T and 3T scanners. Statistical analysis showed a significantly lower mean %CRLB for NAA on the 7T scanner.

No statistically significant difference in %CRLB for NAA was found between the data from both scanners (t(29) = -0.999, p = 0.326, d = -0.182). The mean %CRLB for the 7T scanner was 1.784 (SD = 0.469, Min = 1.000, Max = 2.750), and for the 3T scanner it was 1.904 (SD = 0.581, Min = 1.150, Max = 3.810) (see Figure 48).

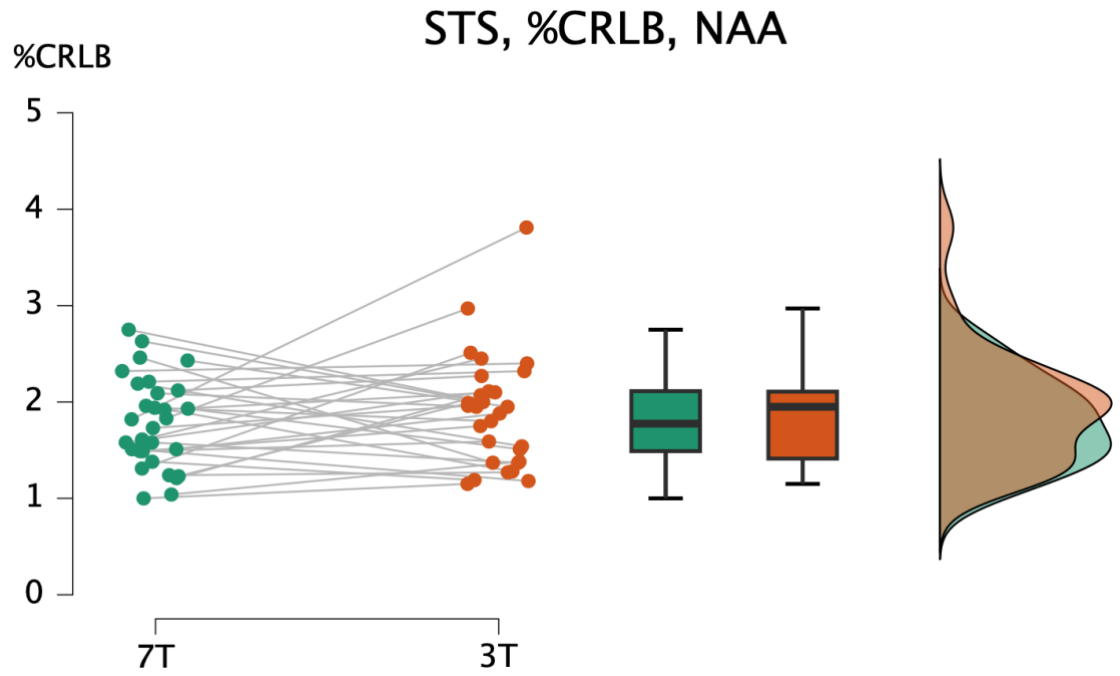

Figure 48. Percentage Cramer-Rao Lower Bound %CRLB comparison for NAA data from the left superior temporal sulcus between the 7T and 3T scanners. No statistically significant difference was observed between the two scanners.

3.3.4. Full width at half maximum (FWHM)

The full width at half maximum (FWHM) parameter describes the width of a single metabolite peak in the spectrum, measured at half of its maximum intensity, in Hertz (Hz). FWHM for data acquired on the 3T scanner was recalculated to compensate for the difference in Larmor frequency. Similar to linewidth, which was calculated for all metabolites together, lower values of FWHM indicate narrower peaks, making them easier to distinguish. FWHM was not calculated correctly for some datasets from the 3T scanner (parameter had a value of 0), likely due to difficulties in separating the glutamate peak. As a result, only the analysis of N-acetylaspartic acid data from the 3T scanner was conducted.

In the medial prefrontal cortex, a statistically significant difference was observed, with narrower peaks for the 3T scanner ($t(33) = 4.027$, $p < 0.001$, $d = 0.691$). The mean FWHM for NAA on the 7T scanner was 13.872 (SD = 3.294, Min = 8.330, Max = 20.440), and for the 3T scanner it was 11.360 (SD = 1.668, Min = 8.700, Max = 16.200) (see Figure 49).

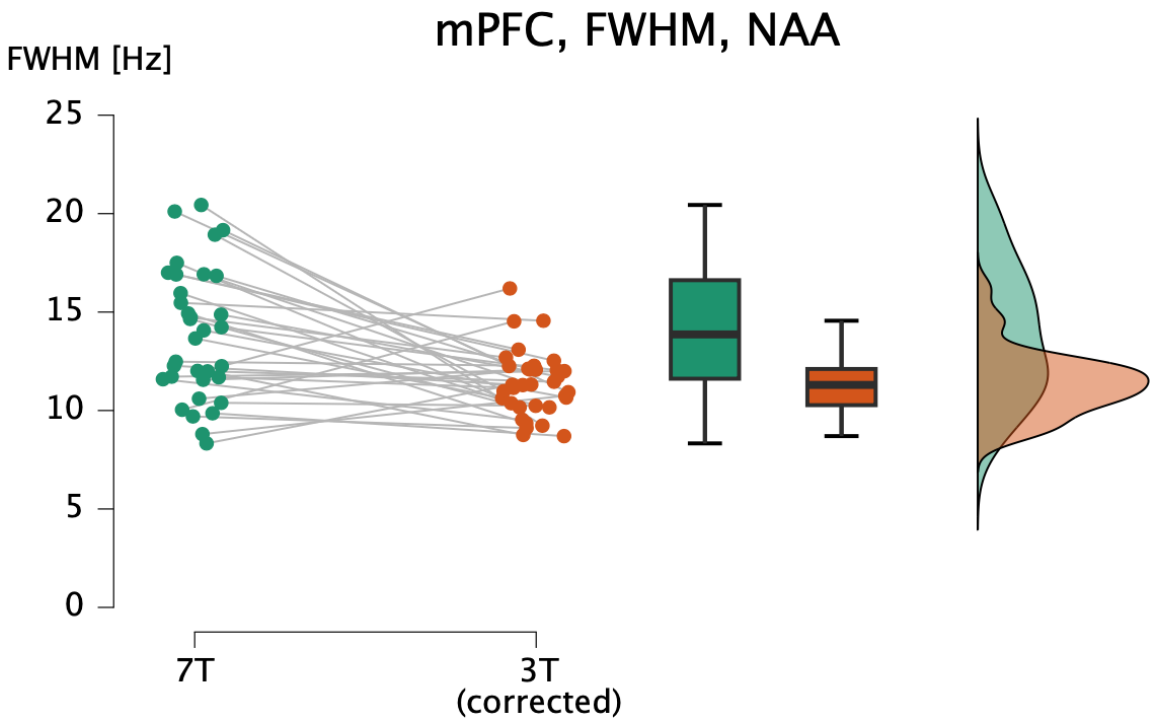

Figure 49. Full width at half maximum (FWHM) comparison for NAA data from the medial prefrontal cortex between the 7T and 3T scanners. Statistical analysis showed a significantly higher mean FWHM for NAA on the 7T scanner.

For the left superior temporal sulcus, no significant difference between scanners was revealed (Wilcoxon signed-rank test: Z = 0.802, p = 0.428, r = 0.168; Shapiro-Wilk test: p = 0.029). The mean FWHM for NAA on the 7T scanner was 13.291 (SD = 3.409, Min = 8.880, Max = 22.550), and for the 3T scanner it was 12.678 (SD = 2.020, Min = 10.010, Max = 18.880) (see Figure 50).

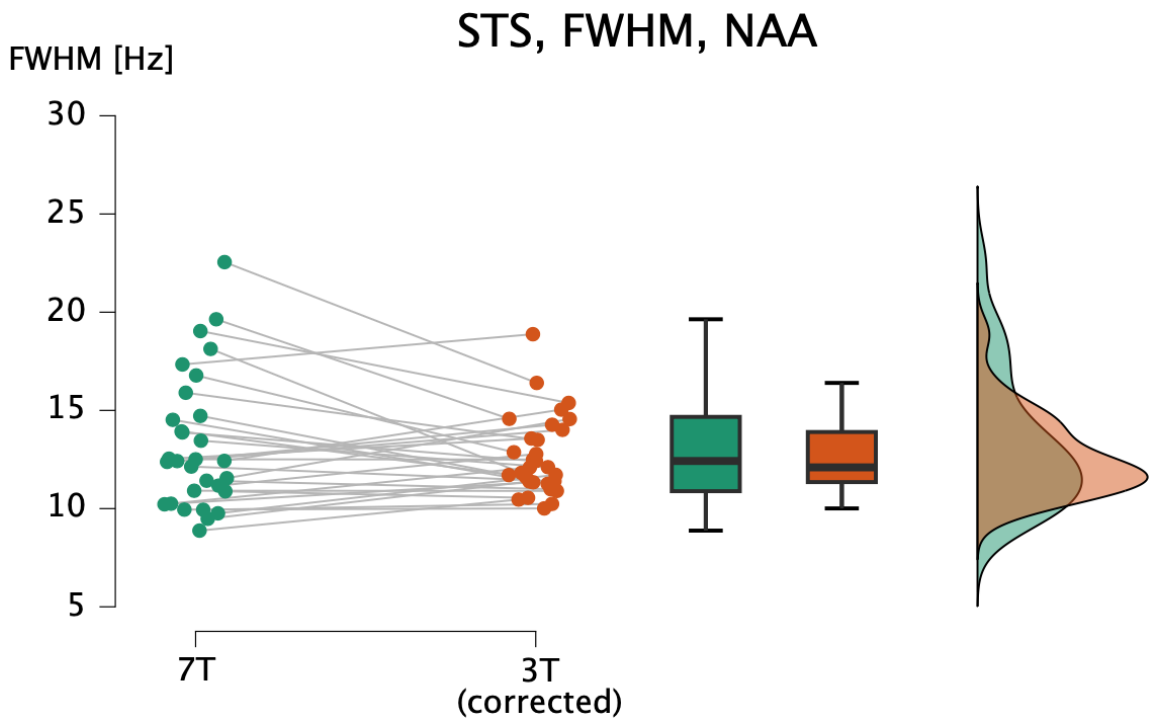

Figure 50. Full width at half maximum (FWHM) comparison for NAA data from the left superior temporal sulcus between the 7T and 3T scanners. No statistically significant difference was observed between the two scanners.

4. Discussion

The present study employed functional magnetic resonance spectroscopy (fMRS) to investigate glutamate modulation in response to reading-related visual stimulation. Regions of interest were localized individually based on fMRI activations rather than anatomical landmarks, allowing us to target reading-sensitive cortex: the left superior temporal sulcus (STS) and the visual word form area (VWFA). In addition, the medial prefrontal cortex (mPFC) was included as an anatomically defined control region. Participants were presented with two types of stimuli: meaningful words in Polish and visually matched false-font strings.

As far as I am aware, this is the first study to examine the main excitatory neurotransmitter during reading tasks within individually defined, functionally localized regions. The sample included individuals with dyslexia and typical readers.

The experimental design addressed five aims. First, we assessed whether glutamate concentrations change in response to reading-related stimulation. Second, we tested the neural noise hypothesis in dyslexia, which proposed a mechanistic account for reading difficulties based on elevated glutamate levels in the STS. Third, we examined regional specificity of the observed effects by testing whether the control region (mPFC) shows any task- or group- related modulation. Fourth, we investigated the temporal dynamics of glutamate response, an aspect still poorly understood. Finally, we compared data quality between 3T and 7T scanners to evaluate methodological implications for future fMRS studies.

In the following sections, we discuss each aim in turn and highlight their implications in the context of prior literature.

4.1. Glutamate concentration changes in response to reading

The first aim of this study was to investigate whether glutamate concentrations change in response to reading-related visual stimulation in two language-sensitive regions: the left superior temporal sulcus (STS) and the visual word form area (VWFA). Both regions were individually defined as showing higher activation for words compared to false font strings in each participant. Based on prior work, we expected a glutamate increase following

language stimulation compared to rest in the STS, which is known to respond to both visual and auditory input (van Atteveldt et al., 2004) and shows higher activation to printed words compared to symbol strings (Dębska et al., 2021). For the VWFA, we anticipated stronger modulation for meaningful words presented in Polish compared to visually matched false-font strings (Dehaene & Cohen, 2011).

For the VWFA, we were unable to obtain sufficient high-quality spectra for reliable analysis. To the best of my knowledge, no previous study has applied MRS or fMRS in this region, most likely due to the technical challenges of voxel placement and spectral acquisition in this anatomically small, susceptibility-prone area near air-filled sinuses (see Figure 51). On the 7T scanner, VWFA spectra were acquired in 42 out of 59 participants, but only 12 met quality criteria. On the 3T scanner, spectra were obtained in 21 out of 40 participants, with only 9 being suitable for analysis. These results highlight the substantial difficulty of reliably measuring neurochemical signals in the VWFA. Future studies could benefit from optimized voxel placement and advanced shimming techniques to improve data quality in this region.

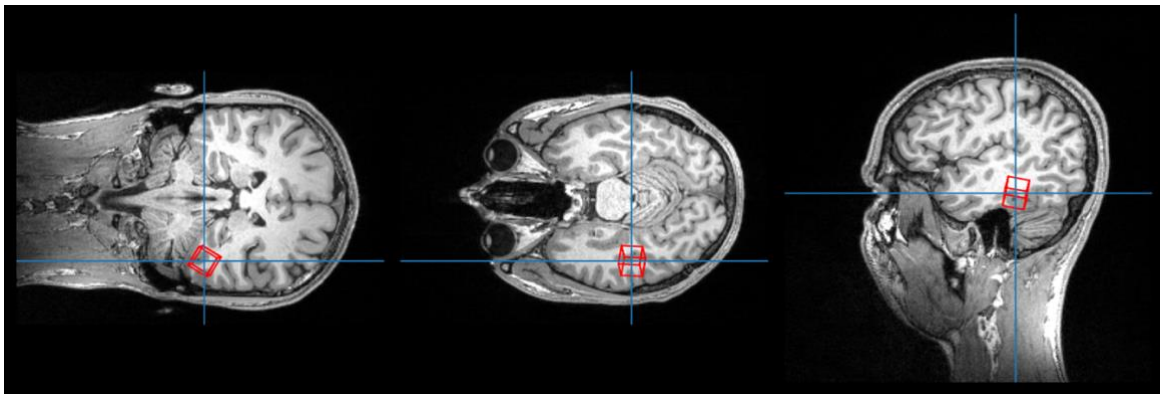

Figure 51. Example of the location of the visual word form area (VWFA) in one participant.

In the STS, no consistent increase in glutamate concentration was observed during linguistic stimulation compared to rest. Although an interaction with sex was initially observed at 7T, it disappeared after BOLD correction, which suggests that the effect was at least partly driven by hemodynamic changes rather than true neurochemical modulation. Interestingly, the direction of the effect differed between sexes: females tended to show higher glutamate levels after both types of visual stimulation (words and false fonts) compared to rest, whereas in males the opposite pattern emerged, but only for word stimuli, where glutamate levels were lower during stimulation than at rest.

Findings between scanners were inconsistent. Results from the 3T scanner showed no significant difference in glutamate concentration between rest and stimulation with words or false font strings. However, after BOLD correction, glutamate levels increased with word stimulation, as expected. A significant interaction between stimulation and sex was observed: females showed a stronger glutamate response to word stimulation compared to both false font strings and rest, while males' glutamate levels remained stable across conditions. Interestingly, females had lower glutamate levels during rest and after false font string stimulation compared to males. Differences between 3T and 7T results may reflect both sample characteristics (a smaller sample was included in the 3T experiment) and the limitations of BOLD correction. These methodological factors should be taken into account when interpreting the observed differences.

Given the established role of the STS in audiovisual integration of letters and speech sounds (Beck et al., 2024) and reading development (Chyl et al., 2018, Feng et al. 2020), the absence of robust glutamate modulation observed here suggests that neurochemical responses in this region may be subtler than expected. To the best of my knowledge, only one previous fMRS study has measured neurometabolite concentrations in the STS. Pasanta et al. (2024), in a small sample of 11 participants, investigated GABA and glutamate responses to social (faces) versus non-social (objects) visual stimuli. They reported no consistent modulation by face stimuli in the STS, while small effects were observed for object stimuli. Overall, the authors concluded that the minimal changes they detected were more likely to reflect basic visual processing than social-specific effects. In contrast, the present study revealed some task-related changes in the STS during reading, potentially reflecting both the different cognitive paradigm and the larger sample size employed here. Importantly, Pasanta et al. (2024) also emphasized that their passive viewing paradigm was not inherently social and might therefore not have been sufficiently engaging to elicit robust neurochemical responses. By analogy, it is possible that the reading task with single short words used in the present study was also not demanding enough to consistently drive measurable metabolite changes in the STS. Indeed, a recent meta-analysis (Turker et al., 2025) of 163 fMRI studies during reading in adults, highlights the role of left STS in sentence processing, with more consistent activation than for single words.

Previous MRS studies have primarily been conducted in the resting state, comparing baseline metabolite concentrations between individuals with and without dyslexia

(e.g., Rae et al., 1998; Pugh et al., 2014). These studies reported group-level differences in glutamate and GABA concentrations in language-related regions, but they did not address dynamic, task-related neurochemical modulation. The present study therefore extends this line of research by applying fMRS in the STS to examine whether glutamate levels change in response to reading-related visual stimuli.

Overall, this study provides important methodological insights and highlights the challenges of probing small cortical regions such as the VWFA, as well as the subtle nature of task-related neurochemical modulation in language-sensitive areas.

4.2. The neural noise hypothesis of dyslexia

The second aim of this study was to test the neuronal noise hypothesis, which proposes that dyslexia may result from elevated glutamate concentrations in the left STS, leading to impaired signal processing (Hancock et al., 2017). Based on this account, we expected group differences in the left STS, with lower glutamate levels in typical readers. Furthermore, the hypothesis predicts that increased glutamate in individuals with dyslexia should correlate with poorer performance on reading-related behavioral tasks such as phonological awareness and rapid automatized naming (RAN).

As expected, participants with dyslexia scored lower on reading-related behavioral tasks, including RAN and phonological awareness at 7T, and a similar pattern was observed at 3T, except for digit and object RAN. These behavioral results are consistent with the previous studies of dyslexia (e.g. Dębska et al., 2022) and provide a basis for testing whether such difficulties are associated with altered neurochemical responses in the STS, as predicted by the neuronal noise hypothesis.

At 7T, no significant group differences were observed in glutamate concentration in the left STS. A trend-level interaction emerged in the mixed dynamic-averaging approach after BOLD correction, with female participants with dyslexia showing elevated glutamate relative to typical readers at specific delays (500 ms after false-font stimulation and 3000 ms after word stimulation). At 3T, a significant group effect was found after BOLD correction when data were grouped by stimulus type, but in the opposite direction of the neuronal noise hypothesis: typical readers showed higher glutamate levels than individuals with dyslexia, irrespective of stimulation type. This effect appeared to be driven primarily

by differences in females as no significant group differences were present in males, and similar trends were observed in the delay-based analysis. Overall, the group effects were inconsistent across scanners and analysis approaches and seemed to depend on participants sex, indicating that the present results do not provide clear support for the neuronal noise hypothesis. At the same time this study highlights the need of including participants sex as an important factor in analyses of the neural basis of dyslexia (Ramus et al., 2018), even though no significant sex effects were observed on behavioral level.

This study represents the first application of fMRS to directly examine the neuronal noise hypothesis in the left STS. As such, there are no prior fMRS findings available for direct comparison, and interpretation must therefore be guided by the broader MRS literature on dyslexia. Previous MRS studies have reported mixed findings regarding the excitatory–inhibitory balance in dyslexia. Some results support the hypothesis, with higher glutamate concentrations in the occipital cortex linked to poorer reading and phonological skills (Pugh et al., 2014; Del Tufo et al., 2018), and reduced glutamatergic metabolites following reading intervention (Cecil et al., 2021). Others, however, point in the opposite direction, such as positive correlations between glutamate and phonological abilities in the anterior cingulate cortex (Lebel et al., 2016), or higher GABA levels in the left inferior frontal gyrus associated with poorer verbal fluency (Nakai & Okanoya, 2016). Adding to this mixed picture, other investigations have reported null findings in anterior cingulate cortex (Horowitz-Kraus et al., 2018) and in temporo-parietal and occipital regions (Kossowski et al., 2019).

In this context, the present findings extend the literature by applying a dynamic, task-based fMRS approach to the STS. The absence of robust group-level differences suggests that the present findings do not provide clear support for the neuronal noise hypothesis in the STS. Although this study employed a novel dynamic fMRS paradigm in a language-related area, the findings remain inconclusive. Future work combining static and dynamic approaches will be essential to determine whether excitatory–inhibitory imbalance contributes to dyslexia.

However, an important limitation of the current MRS literature on dyslexia is the considerable methodological heterogeneity across studies. Investigations have targeted different brain regions, employed diverse acquisition sequences and analysis strategies, and used varying reference approaches for metabolite quantification. Such variability in

methodology substantially complicates cross-study comparisons and may, at least in part, explain the mixed and sometimes contradictory findings regarding glutamate and GABA alterations in dyslexia.

To gain a more comprehensive perspective, it is also useful to consider the evidence from studies using electroencephalography (EEG). Beyond neurochemical studies, complementary insights into the excitatory–inhibitory balance in dyslexia have been obtained from EEG. Turri et al. (2023) reported alterations in oscillatory dynamics, particularly in the beta frequency range, along with changes in the aperiodic component of the spectrum, both of which are thought to reflect disrupted excitatory–inhibitory regulation. Particularly, in the parieto-occipital regions participants with dyslexia presented lower aperiodic exponent and lower beta oscillations linked with higher excitation to inhibition balance as compared to the control group. While EEG does not directly measure glutamate or GABA, such findings provide evidence consistent with the neuronal noise hypothesis. A related multimodal study, which integrated averaged left STS MRS data from the overlapping sample of participants with EEG recordings obtained in an extended sample, also found no support for the neuronal noise hypothesis (Glica et al., 2025). Much larger sample and the convergence across analytic approaches underscores the need to explore other neural mechanisms associated with reading difficulties besides the neural noise hypothesis.

Taken together, these findings indicate that while the neuronal noise hypothesis remains an important theoretical framework, the neurochemical basis of dyslexia is far from conclusive and requires further multimodal, region-specific investigation.

4.3. Regional specificity of fMRS effects

The third aim of this thesis was to examine whether the effects observed in reading-related regions (VWFA and STS) are specific to those areas, or whether similar responses can also be detected in the medial prefrontal cortex (mPFC). Since the mPFC is not directly involved in reading (Richlan et al., 2009), it served as a control region to assess the regional specificity of the detected effects.

The medial prefrontal cortex (mPFC) has been consistently linked to various neurological and psychiatric disorders, including depression, schizophrenia, and autism (Xu et al.,

2019). While it is not traditionally considered a part of the core reading network (Richlan et al., 2009), the mPFC plays a central role in integrating information from other brain regions and translating it into coordinated behavioral output (Xu et al., 2019). It is a critical region for a range of higher-order cognitive processes, such as memory consolidation, rapid learning, and decision-making (Euston et al., 2012). These functions are closely associated with executive functioning, which has been proposed as an important contributor to reading development. Deficits in executive functions, observable even before reading acquisition in early childhood, have been linked to later reading difficulties and may serve as early indicators of dyslexia, offering potential for early intervention (Farah et al., 2021). Furthermore, children with dyslexia who successfully improved their reading and spelling skills showed strong negative intrinsic functional connectivity between the left fusiform gyrus and the right medial prefrontal cortex. This pattern is likely to reflect a compensatory mechanism rather than a normalization of typical connectivity (Koyama et al., 2013).

When considering the response to reading-related stimulation, results in the mPFC revealed a number of interaction effects, though these were inconsistent across scanners and analytic approaches. At 7T, a significant interaction of stimulation type and sex was observed, driven by stronger responses to visual stimulation in females when data were grouped by stimulus type. However, this effect disappeared after BOLD correction. The mixed dynamic-averaging approach also revealed a significant interaction between stimulation type, delay, and group, which remained at a trend level after BOLD correction. At 3T, analyses based on delay indicated significant interactions involving stimulation type with several factors (group, sex, delay, age, and gray matter volume) and these effects remained significant after BOLD correction. A similar interaction pattern was found using the dynamic-averaged approach, again persisting after correction.

In terms of group effects, no main effect of group was detected in the mPFC. While some interactions involving group did emerge, their direction was opposite to the predictions of the neuronal noise hypothesis. Thus, when considered together with the lack of conclusive interaction effects across scanners and analytic approaches, these findings indicate that the mPFC did not exhibit a stable activation pattern in response to reading-related stimulation.

The overall results of the thesis did not reveal a clear and consistent glutamate increase across all examined regions. While some interactions emerged in the STS, these effects were not robust enough to support firm conclusions. In the VWFA, data quality issues

prevented reliable analysis, limiting interpretation for this region. In the mPFC, no systematic pattern of responses was reported. Taken together, these findings suggest that neurochemical responses to reading-related stimulation are likely to be subtle and highly sensitive to methodological factors and group composition. Importantly, the lack of systematic activation in the control region (mPFC) should not be interpreted as an absence of its involvement in reading. Rather, it indicates that the observed variability is unlikely to reflect a global, non-specific cortical response. Instead, the results highlight both the challenges of applying fMRS to study region-specific processes such as reading and the need for methodological improvements in the future work. It is possible that some of the observed responses in the mPFC were partly driven by decision-making demands of the task, as participants were asked to evaluate whether each stimulus was a target or not.

Findings from nearby regions support the idea that prefrontal areas are also capable of showing dynamic neurochemical responses in fMRS. For example, Craven et al. (2024) demonstrated task-related increases in Glx, but not GABA, within the ACC during a cognitive control task, with fMRS-derived BOLD measures correlating with conventional fMRI responses. Similarly, Huang et al. (2015) reported Glx modulation in the mPFC during mental imagery. Together, these results indicate that frontal midline regions can exhibit robust metabolic reactivity under specific cognitive demands. In contrast, the present study did not detect consistent changes in glutamate in the mPFC during reading-related stimulation, suggesting that the engagement of this region may depend on the nature of the recruited cognitive process. Whereas imagery and executive control tasks strongly activate the prefrontal cortex, reading may not impose comparable demands, potentially accounting for the absence of systematic effects in our data.

On the other hand, a recent meta-analysis by Kiemes et al. (2021) emphasized that while GABA levels in the mPFC and ACC show relatively robust and consistent negative associations with local fMRI activation during emotional processing (and similarly in the occipital lobe during visual tasks), evidence for task-related modulation of glutamate is far less consistent. Such effects appear more difficult to capture and often show substantial heterogeneity across studies. From this perspective, the absence of reliable glutamatergic responses in the present study is more likely to reflect methodological challenges in detecting Glu dynamics with fMRS, rather than indicating a true absence of neurochemical involvement of the mPFC.

Another consideration is the anatomical position close to the anterior cingulate cortex (ACC), which has been repeatedly implicated in dyslexia using conventional MRS approaches (see Figure 52). Several studies reported altered metabolite levels in the ACC in children with dyslexia, including associations between glutamate, creatine, and inositol concentrations and phonological processing (Lebel et al., 2016), as well as negative correlations between choline and processing speed in girls with dyslexia (Horowitz-Kraus et al., 2018). More recently, lower concentrations of Glx, Glu, Cr, and NAA in the ACC were linked to better word reading performance in dyslexic children (Cecil et al., 2021). These findings suggest that the ACC may play a role in reading-related difficulties, particularly through its involvement in executive control and attention. Given the close spatial proximity between the ACC and the mPFC, it cannot be excluded that some of the variability observed in the present study partially reflects neurochemical processes similar to those previously reported in the ACC. However, unlike prior studies relying on static MRS measures, the current approach employed dynamic fMRS, which may capture different aspects of glutamatergic function and could explain the lack of direct replication.

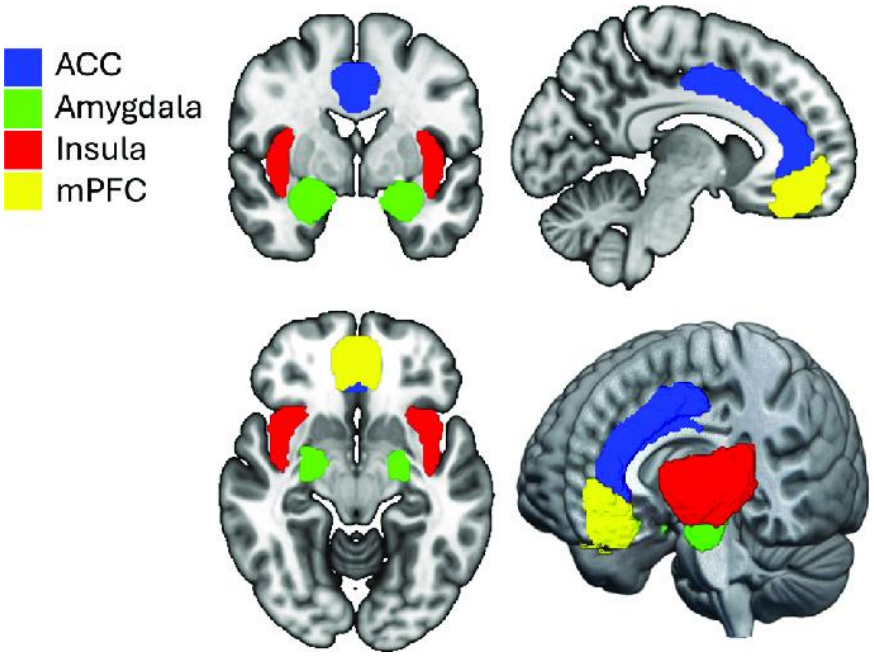

Figure 52. Figure reproduced from Hand et al. (2024). Licensed under CC BY 4.0.

In summary, no reliable glutamatergic responses to reading-related stimulation were detected in the mPFC. However, similarly weak or inconsistent patterns were also found in the reading-related region, the STS, suggesting that the absence of systematic effects in the control region does not reflect a specific lack of involvement of the mPFC in reading.

Instead, it points to the overall subtlety of task-related neurochemical responses under the current experimental conditions.

4.4. The glutamate response function

The fourth aim of this study was to investigate how glutamate concentration changes over time following stimulation. A clear understanding of glutamate temporal dynamics could help in designing fMRS experiments that capture actual neurochemical responses without missing critical changes. To date, only a theoretical response function, derived from a meta-analysis, has been proposed, but it still lacks empirical validation (Mullins, 2018). Based on this theoretical prediction, we acquired data at four time intervals following stimulation onset: 500 ms, 1000 ms, 3000 ms, and 4500 ms. This approach allowed us to sample distinct points along the predicted curve and to investigate glutamate dynamics. Importantly, the stimuli employed in this study were more complex and of longer duration compared to those typically used. For this reason, the selected time intervals included both the early phase, when glutamate levels were expected to peak (500 ms and 1000 ms), as well as the later phase, when a return to baseline was predicted (3000 ms and 4500 ms). Notably, the acquisition at 500 ms occurred during the stimulation. Although additional time points would have provided a more detailed sampling of the glutamate response function, this was not possible due to methodological constraints. Extending the scanning protocol would have risked reduced data quality, increased motion artifacts, and participant fatigue, thereby compromising the validity of the results.

In the present study, the results did not provide reliable evidence for glutamate changes across time points. At 3T, a trend-level effect of delay was observed in the control region, but this effect disappeared after applying BOLD correction, suggesting that it may have been partly driven by vascular rather than neurochemical contributions. Some statistically significant interactions with delay were detected; however, no clear temporal profile of glutamate dynamics emerged. These findings indicate that the shape of the glutamate response function remains difficult to characterize and may depend on several factors, including participants' sex, type of stimulation and magnetic field strength.

An important limitation of the present approach is the relatively small number of transients (MRS signals) averaged within each time window, which reduces the signal-to-noise ratio

of the spectra, as SNR is proportional to the square root of the number of acquisitions (Kreis, 2004). It is therefore possible that subtle effects were hidden in the noise.

From a biological perspective, glutamate changes measured with fMRS may arise from multiple processes, including synaptic release, astrocytic uptake, and metabolic cycling (Pasanta et al., 2023). Fast fluctuations likely reflect neurotransmission, whereas slower dynamics may be linked to clearance and recycling pathways. The coexistence of these mechanisms, occurring on different temporal scales, could partly explain the lack of a single, consistent glutamate response function in the present data. Moreover, recent modelling work suggests that fMRS does not directly capture neurotransmitter synthesis or degradation, but rather reflects shifts of glutamate and GABA between vesicular, extracellular, and cytosolic pools (Lea-Carnal et al., 2023). Such compartmental dynamics provide an additional source of variability that complicates the identification of a clear temporal profile of glutamate.

Taken together, the findings offer only limited support for distinct temporal dynamics of glutamate. The lack of a consistent pattern likely reflects methodological challenges rather than the absence of a true neurochemical response. Future studies with improved temporal resolution and higher signal-to-noise ratio (e.g., through averaging more transients per time point), as well as more engaging stimulation paradigms, will be essential to establish the empirical shape of the glutamate response function. In this light, the present results illustrate both the potential and the challenges of using fMRS to characterize glutamate dynamics.

4.5. fMRS quality and field strength

The last aim of the study was to compare data quality between two different MR scanners: an ultra-high-field 7T system (GE, DISCOVERY 950 MR System) and a standard 3T system (Siemens Trio). Based on theoretical considerations, the 7T scanner is expected to provide higher signal-to-noise ratio (SNR) and better spectral resolution, facilitating more accurate metabolite quantification. Despite these predicted advantages, data acquisition on ultra-high-field scanners involves technical challenges, such as more pronounced susceptibility distortions and higher radiofrequency power deposition, expressed as specific absorption rate (SAR) (Pradhan et al., 2015). Examination of data quality from both

scanners provides insights into whether the anticipated advantages justify these technical difficulties.

Additionally, while 3T scanners are widely available, the 7T scanner used in this study is the only one in Poland and one of the few in this part of Europe. Data quality was assessed using parameters automatically calculated in FSL-MRS for spectra obtained by averaging all 320 transients. These parameters reflect key aspects of spectral quality, such as linewidth, full width at half maximum, signal-to-noise ratio, and Cramer-Rao lower bounds. For linewidth and full width at half maximum (FWHM) of N-acetylaspartate (NAA), values from the 3T system were recalculated to match the Larmor frequency of the 7T system. Signal-to-noise ratio and percentage Cramer-Rao lower bounds (%CRLBs) were compared for NAA and glutamate, a metabolite frequently targeted in fMRS studies.

Linewidth was better for the 3T scanner in the mPFC, while no statistically significant differences were observed in the STS. This suggests that, despite the higher field strength, the ultra-high-field scanner does not necessarily provide narrower linewidths, which may be influenced by technical factors such as a less efficient shimming system of the scanner or local tissue susceptibility distortions. Signal-to-noise ratio (SNR) for glutamate in the mPFC was higher for the 7T scanner, consistent with theoretical expectations of improved sensitivity at ultra-high fields. However, no other statistically significant SNR differences were observed. Notably, for NAA in the mPFC, SNR was higher for the 7T scanner, but this difference did not reach statistical significance. In the left superior temporal sulcus, SNR values were comparable for 3 and 7T for both glutamate and NAA, indicating that the benefit of higher field strength may be region-specific. Regarding %CRLB, which reflects quantification accuracy, data from the 7T scanner were generally better, achieving lower %CRLB for almost all comparisons. The only exception was %CRLB for NAA in the STS, which was not significantly improved at 7T. Interestingly, FWHM of NAA in the mPFC was narrower on the 3T scanner, which is in line with the linewidth results and suggests that, under certain conditions, 3T scanners may still provide competitive spectral resolution if optimally calibrated.

Our results indicate that the advantage of ultra-high-field scanning is region-dependent: while the mPFC benefitted from higher SNR for glutamate at 7T, no such improvement was observed in the STS. This regional variability is in line with findings from Pradhan et al. (2015), who compared three brain regions (ACC, DLPFC, and CSO) and showed that

the superiority of 7T over 3T in terms of SNR and metabolite quantification precision was not uniform across locations. In their study, SNR was consistently higher at 7T, but the magnitude of improvement varied, with the weakest effect observed in a white matter voxel. They also reported lower CRLBs at 7T, which reflects greater accuracy in metabolite quantification. According to the authors, such variability was likely related to technical factors, including the sensitivity profile of phased array coils and stronger B1 inhomogeneity at 7T, which reduces the benefits for deeper brain regions.

Our finding that %CRLB values improved more consistently than linewidth or SNR agrees with earlier field-strength comparisons. Reliable quantification of a broader range of metabolites at 7T, particularly those with weaker or more complex spectral patterns (e.g., Gln, GSH, GABA, NAAG), has been reported, whereas abundant compounds such as NAA, tCr, and tCho showed only limited additional benefit (Terpstra et al., 2016). At the same time, SNR at 7T was found to be nearly twice that at 4T, but linewidths were about 50% broader, resulting in only modest gains in spectral resolution (Tkáč et al., 2009). Similarly, generally higher SNR and improved resolution at 7T, allowing more robust quantification of metabolites such as glutamate, glutamine, and lactate, have also been reported (Mekle et al., 2009). In our own comparison of the 7T and 3T datasets, we observed patterns consistent with these earlier findings. Specifically, glutamate quantification in the medial prefrontal cortex showed clear benefits at 7T, with improved %CRLB and SNR. However, in some cases 3T performed equally well or better, particularly in terms of linewidth and FWHM.

Finally, other studies have emphasized that while 7T provides advantages in terms of SNR and quantification precision, 3T remains a robust and reliable option in longitudinal investigations due to its availability and consistent reproducibility. This perspective is consistent with our findings. While 7T yielded lower %CRLB values for glutamate, 3T nevertheless produced reliable and reproducible spectra that remain valuable in longitudinal investigations (Eftekhari et al., 2025).

Another important factor is that the scanners differed not only in field strength, but also in vendor, which complicates direct comparison: the 7T scanner was made by GE, while the 3T scanner was from Siemens. Although we implemented a standardized, vendor-independent semi-LASER sequence, the hardware and signal processing still differ between the scanners. In addition, the 7T scanner used in this study is experimental, so it

is not fully validated and has limited technical support. By contrast, the 3T scanner is a common, well-known model that has been used for many years and benefits from reliable technical support and a large user community.

Magnetic field homogeneity was improved using advanced first- and second-order shimming algorithms on both scanners. However, the FAMASITO algorithm used on the GE scanner often failed to calculate appropriate shim values. In such cases, only zero- and first-order shims could be applied. This issue did not occur with the 3T scanner, where the FASTMAP algorithm for first- and second-order shimming was available. In theory, better shimming leads to improved magnetic field homogeneity and, consequently, higher quality of the acquired spectra.

Importantly, we assessed the quality of spectra acquired on both the 7T and 3T scanners. To make the comparison fair, the linewidth threshold for 7T spectra (20 Hz) was recalculated for the 3T spectra, resulting in a corresponding threshold of 8.57 Hz. Establishing this common threshold was necessary because linewidth is a critical quality parameter and the scanners have different magnetic field strengths. The recommended linewidth for the 7T scanner is 19 Hz, and all our 7T spectra met this criterion. For the 3T spectra, the expert-recommended threshold is 13 Hz, but in this study we applied stricter criteria (Öz et al., 2021). These higher expectations for the 3T spectra could have influenced other quality parameters. Notably, none of the 3T spectra were excluded from the analysis based solely on linewidth.

The observed differences in data quality suggest that the theoretical advantages of ultra-high-field spectra are not sufficient to recommend performing fMRS experiments exclusively on 7T scanners using the setup employed in this study. Another practical consideration is the higher SAR at 7T, which requires longer repetition times (TR). This, in turn, reduces both the number of signals that can be acquired during a fixed acquisition time and the temporal resolution of the experiment. Furthermore, longer scan durations increase the risk of motion-related artifacts, which can further compromise spectral quality.

Overall, our findings underscore that while 7T provides measurable improvements in metabolite quantification, especially in terms of %CRLB, these benefits are not universal. They depend on both the anatomical region and the metabolite of interest, and must be carefully balanced against the practical challenges of ultra-high-field acquisition.

4.6. Important factors and covariates in fMRS studies

Participants' sex is an important factor, but rarely considered in fMRS studies. From a neurochemical perspective, several studies have reported sex-specific differences in glutamatergic transmission within the prefrontal cortex. In humans, reviews of the literature indicate structural differences in synapses between females and males, with some studies reporting elevated glutamate levels in females (Kniffin & Briand, 2024). In animal models, reduced glutamatergic activity in the mPFC has been found in male mice compared to females (Knouse et al., 2022). Importantly, fMRS evidence indicates that women may show a stronger glutamatergic response to pain stimulation in the anterior cingulate cortex (Archibald et al., 2020), a region in close anatomical proximity to the mPFC. Finally, sex-based differences have also been documented in dyslexia, with evidence suggesting that males and females may present distinct cognitive and neural profiles (Yang et al., 2022; Krafnick & Evans, 2019). Our findings suggest that sex moderates task-evoked glutamate dynamics in both of the examined regions: the mPFC and the left STS. Although some of these effects were reduced after BOLD correction, generally we observed that females showed stronger responses than males to reading related stimulation. Overall these findings suggest that participants' sex should be taken into account in future fMRS analyses, and the results observed in one sex cannot necessarily be generalized to both sexes.

Age-related influences on metabolite concentrations are well established and underline the importance of controlling for age as a covariate in fMRS studies. Developmental studies consistently show that levels of creatine and choline increase from childhood into adulthood, while glutamate exhibits a negative age-related slope in several cortical regions, including occipital and prefrontal areas (Kossowski et al., 2019; Perica et al., 2022). These maturational effects are thought to reflect synaptic pruning (i.e., elimination of weaker or not used synaptic connections), reduced glutamatergic receptor density, and changes in the glutamate–glutamine cycle, as well as the ongoing refinement of excitation–inhibition balance across development (Perica et al., 2022). Moreover, GABA levels follow heterogeneous trajectories, with region-specific decreases during adolescence and early adulthood, suggesting that inhibitory processes also contribute to shifting neurochemical profiles. A recent systematic review highlighted that such developmental variability may partially explain inconsistencies between spectroscopy findings across studies and emphasized age as a potential confounding factor (Kiemes et al., 2021). Although the

current sample comprised participants within a relatively narrow age range, we observed a systematic negative effect of age on glutamate levels in both the mPFC and the left STS. Metabolite ratios were expressed relative to total creatine, which itself is known to increase with age (Kossowski et al., 2019). This approach may have therefore increased susceptibility to age-related bias. Taken together, these findings underscore that differences in participants' age can significantly affect metabolite estimates, and that age should be systematically modeled in future fMRS research to avoid misinterpretation of neurochemical group effects.

Another important confounding factor is the proportion of gray matter within the scanned voxel. Tissue composition can significantly affect metabolite estimates, as Glx concentrations are substantially higher in gray matter than in white matter (McLean et al., 2000). This is further supported by evidence that, in small, well-localized MRS voxels, glutamate levels positively correlate with the gray matter fraction, whereas in larger voxels this relationship is not evident (DeMayo et al., 2023). Here, the effect of gray matter volume was consistently observed in the left STS, which was individually localized and therefore showed higher variation in the proportion of gray matter within the scanned voxel than the mPFC. These observations align with systematic reviews emphasizing that gray matter fraction should be included as a covariate to reduce variability across studies (Kiemes et al., 2021). Biologically, this makes sense, as most neurotransmitter-related metabolites accumulate in gray matter, while white matter and CSF contribute relatively little to the measured signal. These findings collectively highlight the importance of accounting for voxel tissue composition to avoid biased metabolite quantification and misinterpretation of group differences.

4.7. Limitations

The glutamate response measured with fMRS is influenced by stimulation intensity. Ip et al. (2019) demonstrated that in the visual cortex both BOLD and glutamate signals scaled linearly with stimulus contrast, but a significant increase in glutamate concentration was only detected at the highest contrast level. This suggests that high-intensity stimulation may be necessary to elicit measurable neurochemical changes. In line with this, a meta-analysis of fMRS studies reported that the magnitude of glutamate responses varied substantially depending on the stimulation paradigm and modality, with visual stimulation typically yielding small effects (~2.3%), whereas stimuli such as pain could induce much larger

changes (~14.5%) (Mullins, 2018). These findings indicate that both the type and intensity of stimulation are critical determinants of the detectability of glutamate changes in fMRS. Perhaps, reading single words, even tough being a salient stimulus was not “intensive” enough for adolescents and adults and we should have employed sentenced instead.

The temporal resolution of the paradigm may also influence the detectability of neurochemical responses. Long stimulation and rest blocks can blur transient fluctuations, whereas short blocks may fail to produce measurable changes. Here we opted for an event related paradigm, which short stimulation of 850 ms and trials from both conditions: words and false fonts intermixed within a stimulation block. Possibly, longer stimulation or block design could have produced larger glutamate response. Moreover, individual differences in reaction times and attentional engagement introduce temporal variability, which can further reduce the sensitivity of both averaged and dynamic analyses. These design-related factors should therefore be carefully considered in future fMRS studies.

Results of this project showed some differences between the experiment conducted on the 7T scanner and the one on the 3T scanner. These differences may partly arise from sample size. Most participants scanned at the lower-field scanner were also examined at 7T, but some additional participants took part only in the 3T experiment. In contrast, the 7T study was carried out on an extended sample, which could have influenced the results. The relatively small subgroup sizes (after dividing by group and sex) may have reduced statistical power and could have made comparisons more difficult. Additionally, at 3T glutamate and glutamine signals often overlap, whereas at 7T these metabolites can be more reliably distinguished. This methodological difference may complicate direct comparisons between field strengths. However, in the present project, strict quality criteria were applied to 3T spectra, linewidths were directly compared between 3T and 7T, which allowed glutamate estimates to be considered reliable. Another important factor is the reproducibility of fMRS experiments. The recently proposed Big fMRS Project aims to investigate the sources of variability in fMRS results by collecting data across multiple laboratories worldwide, using different scanners and participant samples, but applying exactly the same paradigm and acquisition parameters. This large-scale effort is expected to shed light on why results may differ even when the same paradigm is used, and to provide a basis for establishing consensus guidelines on how to acquire reliable dynamic MRS data.

Recently, dynamic approaches in fMRS analysis, where a common model is used for all transients simultaneously without averaging, have become popular (Clark et al., 2024, Craven et al., 2024). This method, based on a general linear model with a design matrix, allows researchers to model expected responses as well as confounding effects such as BOLD contributions or motion artefacts. A major advantage of dynamic analysis is the ability to better capture metabolite changes during stimulation and to more effectively account for possible vascular influences. However, in the present study the application of such an approach was highly challenging due to the complexity of the experimental design, which involved eight different types of MRS acquisitions. The scanning sessions were also very long (over 2 hours in total), making motion artefacts a serious concern. To address this, frequency shifts were extracted for each voxel and included as regressors in the design matrix. However, this approach has important limitations, as sudden head movements are not restricted to single transients but can shift the head to a new position, thereby affecting subsequent signals as well. There is no evidence that the frequency shift is able to capture variability in concentration related to subject's movements within the model. A potentially more accurate strategy would be to model each motion event ("unlike" signal) with a dedicated column in the design matrix. However, it remains unclear how to identify such unreliable data without an additional navigator scan or external measurement. When we attempted dynamic modeling on the 7T data, two issues raised concerns about the validity of the results. First, the estimated glutamate time courses appeared to closely follow one of the proposed regressors in a random manner, which seemed unlikely given the low SNR of single transients and suggested possible overfitting. Second, the dynamic analysis indicated lower glutamate concentrations during stimulation compared to rest, which was not consistent with the results obtained from the averaged analysis, in which the number of signals per group was equalized. This discrepancy could partly be explained by the long rest periods in the paradigm, leading to a higher number of transients acquired during rest compared to other groups. Averaging across participants might mitigate some of these problems, but individual variability in timing and susceptibility to motion artefacts could mask true effects. Future analyses could benefit from applying the Metropolis–Hastings (MH) algorithm, which provides a more robust iterative fitting framework.

Direct modeling of the BOLD effect was not possible in this study, as no fMRI signal was acquired simultaneously with fMRS, and unsuppressed water scans interleaved with the functional task were not collected. To compensate for this limitation, a correction procedure

based on spectral broadening using the Cr FWHM was applied. This approach worked reliably for the ultra-high-field data, but for the 3T scanner some spectra could not be corrected due to technical reasons. For those spectra, we also tested NAA-based broadening, as NAA is typically the most prominent singlet in the spectra, but this approach did not resolve the problem either and failed to correct for the BOLD effect. Therefore, while the broadening procedure was designed to minimize vascular confounds, it cannot be excluded that it was ineffective or even introduced additional spectral distortions. To account for the BOLD effect before data averaging, it is advisable to measure it using water or an additional functional scan. Dynamic modeling would also benefit from the additional information about line broadening that is not directly related to changes in concentration.

Voxel size in this experiment was 15 × 15 × 15 mm^3, which can be considered relatively small for fMRS studies. The choice of a smaller voxel involves a trade-off. On the one hand, it reduces the SNR, which increases with larger voxel volumes. On the other hand, a smaller voxel improves shimming, resulting in narrower linewidths and partially compensating for the SNR loss (Wilson et al., 2019). Moreover, it allows for more accurate placement within gray matter. Importantly, the relatively small voxel enabled positioning in anatomically challenging regions such as the STS and the VWFA, minimizing lipid contamination from the skull and ensuring localization in functionally defined areas based on fMRI localizers.

To minimize the impact of between-subject variability while focusing specifically on glutamate dynamics and maintaining consistency with the dynamic analysis, an averaged tCr reference across all time points was used (calculated separately for data grouped by stimulation type and by delay). This approach also facilitated direct comparison between the averaged and mixed dynamic–averaged analyses, where tCr was treated as a fixed parameter throughout the experiment. Nevertheless, using percentage signal change or an interleaved water reference would allow analysis of glutamate changes independent of other reference metabolites.

A further limitation concerns the composition of the basis sets used for quantification. There is ongoing debate on how many and which metabolites should be included. Too few metabolites can result in higher residuals and misattribution of signals, whereas very large basis sets may increase the risk of splitting contributions between overlapping metabolites. This challenge has been emphasized previously (Hofmann et al., 2002). In our study, the

basis set for the 7T scanner was more extensive (27 metabolites) than for the 3T scanner (18 metabolites). Although 7T did not consistently provide narrower linewidths or higher SNR across regions, %CRLB values were generally lower, indicating more reliable quantification. This supports the use of a more complex basis set at 7T, since omitting relevant metabolites at this field strength could have led to systematic fitting errors and to signals being wrongly assigned to other metabolites. Nevertheless, the difference in basis set size between scanners remains a potential source of variability when interpreting cross-field comparisons.

Tests conducted on the 3T scanner demonstrated that a TR of 3 seconds was sufficient to acquire good-quality fMRS data, while maintaining energy deposition within permissible limits. However, when the experiment was transferred to the 7T scanner, test runs revealed that for some participants the SAR exceeded safety thresholds, which caused the scanner to shut down despite the use of a semi-LASER sequence. Consequently, the repetition time had to be increased to achieve acceptable SAR values. This adjustment required a substantial modification of the original experimental paradigm, which had been designed to acquire 400 transients within a 20-minute acquisition. The longer TR reduced the total number of acquired spectra, thereby limiting the number of transients available for averaging at the group level.

Despite the growing interest in fMRS, there are currently no expert consensus guidelines specifically addressing fMRS methodology. In recent years, several consensus documents have been published for standard MRS, including recommendations on motion artifact correction (Andronesi et al., 2021), magnetic resonance spectroscopic imaging (MRSI; Maudsley et al., 2021), edited MRS protocols for metabolites like GABA (Choi et al., 2021), and reporting standards for MRS studies (Lin et al., 2021). However, none of these documents provide dedicated guidance for fMRS. Such guidelines would be highly beneficial for researchers navigating this complex field, offering support in areas such as managing BOLD-related confounds, selecting appropriate paradigm designs, optimizing acquisition timing based on metabolite dynamics, establishing quality control thresholds, and determining best practices for data analysis.

In sum, the study underscores that detecting task-related glutamate changes in reading is challenging and strongly shaped by region, sex, field strength, and analytic strategy. Future progress will require standardized fMRS protocols, better BOLD correction, and

integration with other neuroimaging methods to clarify the neurochemical basis of reading and dyslexia.

5. Summary and conclusions

To sum up the findings described in this thesis:

- Task-related glutamate changes after reading stimuli were observed mainly in females, but these effects did not remain after BOLD correction.
- No group differences were detected in the superior temporal sulcus, providing no support for the neuronal noise hypothesis.
- Glutamate responses were not restricted to reading-related areas, but also in medial prefrontal cortex, underscoring the importance of including both task-relevant and control regions in fMRS studies.
- Due to variability across sex, group, and stimulus type, a consistent glutamate response function could not be defined, likely because of insufficient data quality at analyzed time points.
- Differences between 7T and 3T scanners were observed mainly in %CRLB values, while linewidth in the mPFC was better at 3T.
- Sex-related interactions were evident, emphasizing the need to consider sex as a between-subject factor.
- Age significantly influenced glutamate levels, highlighting the importance of controlling for developmental factors in future work.
- Low data quality in the VWFA led to its exclusion, despite its high relevance for reading and potential to provide important insights.

References

1. Andrade, C. S., Otaduy, M. C. G., Park, E. J., & Leite, C. C. (2014). Phosphorus-31 MR spectroscopy of the human brain: Technical aspects and biomedical applications. *International Journal of Current Research and Review, 6*(9), 41–57.
2. Andronesi, O. C., Bhattacharyya, P. K., Bogner, W., Choi, I.-Y., Hess, A. T., Lee, P., Meintjes, E. M., Tisdall, M. D., Zaitsev, M., & van der Kouwe, A. (2021). Motion correction methods for MRS: Experts' consensus recommendations. *NMR in Biomedicine, 34*(5), e4364. https://doi.org/10.1002/nbm.4364
3. Apšvalka, D., Gadie, A., Clemence, M., & Mullins, P. G. (2015). Event-related dynamics of glutamate and BOLD effects measured using functional magnetic resonance spectroscopy (fMRS) at 3T in a repetition suppression paradigm. *NeuroImage*, 118, 292–300. https://doi.org/10.1016/j.neuroimage.2015.06.015
4. Archibald, J., MacMillan, E. L., Graf, C., Kozlowski, P., Laule, C., & Kramer, J. L. K. (2020). Metabolite activity in the anterior cingulate cortex during a painful stimulus using functional MRS. *Scientific reports*, *10*(1), 19218. https://doi.org/10.1038/s41598-020-76263-3
5. Arnett, A. B., Pennington, B. F., Peterson, R. L., Willcutt, E. G., DeFries, J. C., & Olson, R. K. (2017). Explaining the sex difference in dyslexia. *Journal of child psychology and psychiatry, and allied disciplines*, *58*(6), 719–727. https://doi.org/10.1111/jcpp.12691
6. Awramiuk, E., & Krasowicz-Kupis, G. (2014). Reading and spelling acquisition in Polish: Educational and linguistic determinants. L1-Educational Studies in Language and Literature, 14(2), 1-24. https://doi.org/10.17239/L1ESLL-2014.01.13
7. Baltzer, P. A., & Dietzel, M. (2013). Breast lesions: diagnosis by using proton MR spectroscopy at 1.5 and 3.0 T–systematic review and meta-analysis. *Radiology*, *267*(3), 735–746. https://doi.org/10.1148/radiol.13121856
8. Beck, J., Chyl, K., Dębska, A., Łuniewska, M., van Atteveldt, N., & Jednoróg, K. (2024). Letter-speech sound integration in typical reading development during the first years of formal education. *Child development*, *95*(4), e236–e252. https://doi.org/10.1111/cdev.14080

9. Begley, J. K., Redpath, T. W., Bolan, P. J., & Gilbert, F. J. (2012). In vivo proton magnetic resonance spectroscopy of breast cancer: a review of the literature. *Breast cancer research : BCR*, *14*(2), 207. https://doi.org/10.1186/bcr3132
10. Bellomo, G., Marcocci, F., Bianchini, D., Mezzenga, E., D'Errico, V., Menghi, E., Zannoli, R., & Sarnelli, A. (2016). MR spectroscopy in prostate cancer: New algorithms to optimize metabolite quantification. *PLOS ONE, 11*(11), e0165730. https://doi.org/10.1371/journal.pone.0165730
11. Bieder, A., Yoshihara, M., Katayama, S., Krjutškov, K., Falk, A., Kere, J., & Tapia-Páez, I. (2020). Dyslexia Candidate Gene and Ciliary Gene Expression Dynamics During Human Neuronal Differentiation. *Molecular neurobiology*, *57*(7), 2944–2958. https://doi.org/10.1007/s12035-020-01905-6
12. Bogdanowicz, M., Sajewicz-Radtke, U., Radtke, B. M., Kalka, D., Karpińska, E., Bogdanowicz, K. M., & Łockiewicz, M. (2016). Bateria metod diagnozy przyczyn niepowodzeń szkolnych u uczniów powyżej 16. roku życia. BATERIA-16 plus [B-16]. [A battery of methods for diagnosing the causes of school failure in students over 16 years of age. BATTERY-16 plus [B-16]]. Pracownia Testów Psychologicznych i Pedagogicznych [Laboratory of Psychological and Pedagogical Tests]: Gdańsk, Poland.
13. Boska, M. D., Dash, P. K., Knibbe, J., Epstein, A. A., Akhter, S. P., Fields, N., High, R., Makarov, E., Bonasera, S., Gelbard, H. A., Poluektova, L. Y., Gendelman, H. E., & Gorantla, S. (2014). Associations between brain microstructures, metabolites, and cognitive deficits during chronic HIV-1 infection of humanized mice. *Molecular neurodegeneration*, *9*, 58. https://doi.org/10.1186/1750-1326-9-58
14. Brem, S., Maurer, U., Kronbichler, M., Schurz, M., Richlan, F., Blau, V., Reithler, J., van der Mark, S., Schulz, E., Bucher, K., Moll, K., Landerl, K., Martin, E., Goebel, R., Schulte-Körne, G., Blomert, L., Wimmer, H., & Brandeis, D. (2020). Visual word form processing deficits driven by severity of reading impairments in children with developmental dyslexia. *Scientific reports*, *10*(1), 18728. https://doi.org/10.1038/s41598-020-75111-8
15. Bruno, J. L., Lu, Z. L., & Manis, F. R. (2013). Phonological processing is uniquely associated with neuro-metabolic concentration. *NeuroImage*, *67*, 175–181. https://doi.org/10.1016/j.neuroimage.2012.10.092

16. Buxton, R. B. (2009). *Introduction to functional magnetic resonance imaging: Principles and techniques* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511605505
17. Cecil, K. M., Brunst, K. J., & Horowitz-Kraus, T. (2021). Greater reading gain following intervention is associated with low magnetic resonance spectroscopy derived concentrations in the anterior cingulate cortex in children with dyslexia. *Brain research*, *1759*, 147386. https://doi.org/10.1016/j.brainres.2021.147386
18. Chavhan, G. B., Babyn, P. S., Thomas, B., Shroff, M. M., & Haacke, E. M. (2009). Principles, techniques, and applications of T2*-based MR imaging and its special applications. *RadioGraphics, 29*(5), 1433–1449. https://doi.org/10.1148/rg.295095034
19. Chen, L., Wassermann, D., Abrams, D. A., Kochalka, J., Gallardo-Diez, G., & Menon, V. (2019). The visual word form area (VWFA) is part of both language and attention circuitry. *Nature communications*, *10*(1), 5601. https://doi.org/10.1038/s41467-019-13634-z
20. Chen, X., Li, J., Chen, D., Zhou, Y., Tu, Z., Lin, M., Kang, T., Lin, J., Gong, T., Zhu, L., Zhou, J., Lin, O., Guo, J., Dong, J., Guo, D., & Qu, X. (2024). CloudBrain-MRS: An intelligent cloud computing platform for in vivo magnetic resonance spectroscopy preprocessing, quantification, and analysis. *Journal of Magnetic Resonance, 358*, 107601. https://doi.org/10.1016/j.jmr.2023.107601
21. Choi, I. Y., Andronesi, O. C., Barker, P., Bogner, W., Edden, R. A. E., Kaiser, L. G., Lee, P., Marjańska, M., Terpstra, M., & de Graaf, R. A. (2021). Spectral editing in ^{1}H magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR in biomedicine*, *34*(5), e4411. https://doi.org/10.1002/nbm.4411
22. Chyl, K., Fraga-González, G., Brem, S., & Jednoróg, K. (2021). Brain dynamics of (a)typical reading development-a review of longitudinal studies. *NPJ science of learning*, *6*(1), 4. https://doi.org/10.1038/s41539-020-00081-5

23. Chyl, K., Kossowski, B., Dębska, A., Łuniewska, M., Banaszkiewicz, A., Żelechowska, A., Frost, S. J., Mencl, W. E., Wypych, M., Marchewka, A., Pugh, K. R., & Jednoróg, K. (2018). Prereader to beginning reader: changes induced by reading acquisition in print and speech brain networks. *Journal of child psychology and psychiatry, and allied disciplines*, *59*(1), 76–87. https://doi.org/10.1111/jcpp.12774

24. Clarke, W. T., Bell, T. K., Emir, U. E., Mikkelsen, M., Oeltzschner, G., Shamaei, A., Soher, B. J., & Wilson, M. (2022). NIfTI-MRS: A standard data format for magnetic resonance spectroscopy. *Magnetic resonance in medicine*, *88*(6), 2358–2370. https://doi.org/10.1002/mrm.29418

25. Clarke, W. T., Ligneul, C., Cottaar, M., Ip, I. B., & Jbabdi, S. (2024). Universal dynamic fitting of magnetic resonance spectroscopy. *Magnetic resonance in medicine*, *91*(6), 2229–2246. https://doi.org/10.1002/mrm.30001

26. Clarke, W. T., Stagg, C. J., & Jbabdi, S. (2021). FSL-MRS: An end-to-end spectroscopy analysis package. *Magnetic resonance in medicine*, *85*(6), 2950–2964. https://doi.org/10.1002/mrm.28630

27. Craven, A. R., Dwyer, G., Ersland, L., Kazimierczak, K., Noeske, R., Sandøy, L. B., Johnsen, E., & Hugdahl, K. (2024). GABA, glutamatergic dynamics and BOLD contrast assessed concurrently using functional MRS during a cognitive task. *NMR in biomedicine*, *37*(3), e5065. https://doi.org/10.1002/nbm.5065

28. Cudalbu, C., Behar, K. L., Bhattacharyya, P. K., Bogner, W., Borbath, T., de Graaf, R. A., Gruetter, R., Henning, A., Juchem, C., Kreis, R., Lee, P., Lei, H., Marjańska, M., Mekle, R., Murali-Manohar, S., Považan, M., Rackayová, V., Simicic, D., Slotboom, J., … Mlynárik, V. (2021). Contribution of macromolecules to brain ^{1}H MR spectra: Experts' consensus recommendations. *NMR in Biomedicine, 34*(5), e4393. https://doi.org/10.1002/nbm.4393

29. Dębska, A., Banfi, C., Chyl, K., Dzięgiel-Fivet, G., Kacprzak, A., Łuniewska, M., Plewko, J., Grabowska, A., Landerl, K., & Jednoróg, K. (2021). Neural patterns of word processing differ in children with dyslexia and isolated spelling deficit. *Brain structure & function*, *226*(5), 1467–1478. https://doi.org/10.1007/s00429-021-02255-2

30. Dębska, A., Łuniewska, M., Zubek, J., Chyl, K., Dynak, A., Dzięgiel-Fivet, G., Plewko, J., Jednoróg, K., & Grabowska, A. (2022). The cognitive basis of dyslexia in school-aged children: A multiple case study in a transparent orthography. *Developmental science*, *25*(2), e13173. https://doi.org/10.1111/desc.13173
31. Deelchand, D. K., Berrington, A., Noeske, R., Joers, J. M., Arani, A., Gillen, J., Schär, M., Nielsen, J.-F., Peltier, S., Seraji-Bozorgzad, N., Landheer, K., Juchem, C., Soher, B. J., Noll, D. C., Kantarci, K., Ratai, E. M., Mareci, T. H., Barker, P. B., & Öz, G. (2021). Across-vendor standardization of semi-LASER for single-voxel MRS at 3T. *NMR in Biomedicine, 34*(5), e4218**.** https://doi.org/10.1002/nbm.4218
32. Deelchand, D. K., Berrington, A., Noeske, R., Joers, J. M., Arani, A., Gillen, J., Schär, M., Nielsen, J. F., Peltier, S., Seraji-Bozorgzad, N., Landheer, K., Juchem, C., Soher, B. J., Noll, D. C., Kantarci, K., Ratai, E. M., Mareci, T. H., Barker, P. B., & Öz, G. (2021). Across-vendor standardization of semi-LASER for single-voxel MRS at 3T. *NMR in biomedicine*, *34*(5), e4218. https://doi.org/10.1002/nbm.4218
33. Dehaene, S. (2009). *Reading in the Brain: The Science and Evolution of a Human Invention.* New York: Viking Penguin.
34. Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in cognitive sciences*, *15*(6), 254–262. https://doi.org/10.1016/j.tics.2011.04.003
35. Del Tufo, S. N., Frost, S. J., Hoeft, F., Cutting, L. E., Molfese, P. J., Mason, G. F., Rothman, D. L., Fulbright, R. K., & Pugh, K. R. (2018). Neurochemistry Predicts Convergence of Written and Spoken Language: A Proton Magnetic Resonance Spectroscopy Study of Cross-Modal Language Integration. *Frontiers in psychology*, *9*, 1507. https://doi.org/10.3389/fpsyg.2018.01507
36. DeMayo, M. M., McGirr, A., Selby, B., MacMaster, F. P., Debert, C. T., & Harris, A. D. (2023). Consistency of frontal cortex metabolites quantified by magnetic resonance spectroscopy within overlapping small and large voxels. *Scientific reports*, *13*(1), 2246. https://doi.org/10.1038/s41598-023-29190-y
37. Deshmukh, S., Subhawong, T., Carrino, J. A., & Fayad, L. (2014). Role of MR spectroscopy in musculoskeletal imaging. *The Indian journal of radiology & imaging*, *24*(3), 210–216. https://doi.org/10.4103/0971-3026.137024

38. Dienel G. A. (2012). Brain lactate metabolism: the discoveries and the controversies. *Journal of Cerebral Blood Flow & Metabolism*, *32*(7), 1107–1138. https://doi.org/10.1038/jcbfm.2011.175

39. 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*(10), e76823. https://doi.org/10.1371/journal.pone.0076823

40. Duarte, J. M. N., & Xin, L. (2019). Magnetic resonance spectroscopy in schizophrenia: Evidence for glutamatergic dysfunction and impaired energy metabolism. *Neurochemical Research, 44*(1), 102–116. https://doi.org/10.1007/s11064-018-2521-z

41. Edden, R. A., Puts, N. A., Harris, A. D., Barker, P. B., & Evans, C. J. (2014). Gannet: A batch-processing tool for the quantitative analysis of gamma-aminobutyric acid–edited MR spectroscopy spectra. *Journal of magnetic resonance imaging : JMRI*, *40*(6), 1445–1452. https://doi.org/10.1002/jmri.24478

42. Eftekhari, Z., Shaw, T. B., Deelchand, D. K., Marjańska, M., Bogner, W., & Barth, M. (2025). Reliability and Reproducibility of Metabolite Quantification Using ^{1}H MRS in the Human Brain at 3 T and 7 T. *NMR in biomedicine*, *38*(8), e70087. https://doi.org/10.1002/nbm.70087

43. Euston, D. R., Gruber, A. J., & McNaughton, B. L. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, *76*(6), 1057–1070. https://doi.org/10.1016/j.neuron.2012.12.002

44. Evans, T. M., Flowers, D. L., Napoliello, E. M., & Eden, G. F. (2014). Sex-specific gray matter volume differences in females with developmental dyslexia. *Brain structure & function*, *219*(3), 1041–1054. https://doi.org/10.1007/s00429-013-0552-4

45. Faghihi, R., Zeinali-Rafsanjani, B., Mosleh-Shirazi, M.-A., Saeedi-Moghadam, M., Lotfi, M., Jalli, R., & Iravani, V. (2017). Magnetic resonance spectroscopy and its clinical applications: A review. *Journal of Medical Imaging and Radiation Sciences, 48*(3), 233–253. https://doi.org/10.1016/j.jmir.2017.06.004

46. Farah, R., Ionta, S., & Horowitz-Kraus, T. (2021). Neuro-Behavioral Correlates of Executive Dysfunctions in Dyslexia Over Development From Childhood to Adulthood. *Frontiers in psychology*, *12*, 708863. https://doi.org/10.3389/fpsyg.2021.708863

47. Fecenec, D., Jaworowska, A., Matczak, A., Stańczak, J., & Zalewska, E. (2013). Test szybkiego nazywania (TSN) [Rapid Automatized Naming Task]. Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego [Laboratory of Psychological Tests of the Polish Psychological Association]: Warszawa, Poland.
48. Feng, X., Altarelli, I., Monzalvo, K., Ding, G., Ramus, F., Shu, H., Dehaene, S., Meng, X., & Dehaene-Lambertz, G. (2020). A universal reading network and its modulation by writing system and reading ability in French and Chinese children. *eLife*, *9*, e54591. https://doi.org/10.7554/eLife.54591
49. Fisher, S. E., & Francks, C. (2006). Genes, cognition and dyslexia: learning to read the genome. *Trends in cognitive sciences*, *10*(6), 250–257. https://doi.org/10.1016/j.tics.2006.04.003
50. Ford, T. C., & Crewther, D. P. (2016). A comprehensive review of the 1H-MRS metabolite spectrum in autism spectrum disorder. *Frontiers in Molecular Neuroscience, 9*, 14. https://doi.org/10.3389/fnmol.2016.00014
51. Gajdošík, M., Landheer, K., Swanberg, K. M., & Juchem, C. (2021). INSPECTOR: free software for magnetic resonance spectroscopy data inspection, processing, simulation and analysis. *Scientific reports*, *11*(1), 2094. https://doi.org/10.1038/s41598-021-81193-9
52. Garwood, M., & DelaBarre, L. (2001). The return of the frequency sweep: Designing adiabatic pulses for contemporary NMR. *Journal of Magnetic Resonance, 153*(2), 155–177. https://doi.org/10.1006/jmre.2001.2340
53. Georgiewa, P., Rzanny, R., Gaser, C., Gerhard, U. J., Vieweg, U., Freesmeyer, D., Mentzel, H. J., Kaiser, W. A., & Blanz, B. (2002). Phonological processing in dyslexic children: a study combining functional imaging and event related potentials. *Neuroscience letters*, *318*(1), 5–8. https://doi.org/10.1016/s0304-3940(01)02236-4
54. Giacometti, L. L., & Barker, J. M. (2020). Sex differences in the glutamate system: Implications for addiction. *Neuroscience and biobehavioral reviews*, *113*, 157–168. https://doi.org/10.1016/j.neubiorev.2020.03.010

55. Gill, S. K., Rose, H. E. L., Wilson, M., Rodriguez Gutierrez, D., Worthington, L., Davies, N. P., MacPherson, L., Hargrave, D. R., Saunders, D. E., Clark, C. A., Payne, G. S., Leach, M. O., Howe, F. A., Auer, D. P., Jaspan, T., Morgan, P. S., Grundy, R. G., Avula, S., Pizer, B., Arvanitis, T. N., & Peet, A. C. (2024). Characterisation of paediatric brain tumours by their MRS metabolite profiles. *NMR in Biomedicine, 37*(5), e5101. https://doi.org/10.1002/nbm.5101

56. Glica, A., Wasilewska, K., Jurkowska, J., Żygierewicz, J., Kossowski, B., & Jednoróg, K. (2025). Reevaluating the neural noise in dyslexia using biomarkers from electroencephalography and high-resolution magnetic resonance spectroscopy. *eLife, 13,* RP99920. https://doi.org/10.7554/eLife.99920

57. Glica, A., Wasilewska, K., Kossowski, B., Żygierewicz, J., & Jednoróg, K. (2024). Sex Differences in Low-Level Multisensory Integration in Developmental Dyslexia. *The Journal of Neuroscience*, *44*(3), e0944232023. https://doi.org/10.1523/JNEUROSCI.0944-23.2023

58. Gonçalves-Ribeiro, J., Pina, C. C., Sebastião, A. M., & Vaz, S. H. (2019). Glutamate Transporters in Hippocampal LTD/LTP: Not Just Prevention of Excitotoxicity. *Frontiers in cellular neuroscience*, *13*, 357. https://doi.org/10.3389/fncel.2019.00357

59. Govindaraju, V., Young, K., & Maudsley, A. A. (2000). Proton NMR chemical shifts and coupling constants for brain metabolites. *NMR in Biomedicine, 13*(3), 129–153. https://doi.org/10.1002/1099-1492(200005)13:3

60. Gruetter, R. (1993). Automatic, localized in vivo adjustment of all first- and second-order shim coils. *Magnetic Resonance in Medicine, 29*(6), 804–811. https://doi.org/10.1002/mrm.1910290613

61. Gruetter, R., Adriany, G., Choi, I.-Y., Henry, P.-G., Lei, H., & Öz, G. (2003). Localized in vivo ^{13}C NMR spectroscopy of the brain. *NMR in Biomedicine, 16*(6–7), 313–338. https://doi.org/10.1002/nbm.841

62. Gruetter, R., Weisdorf, S. A., Rajanayagan, V., Terpstra, M., Merkle, H., Truwit, C. L., Garwood, M., Nyberg, S. L., & Uğurbil, K. (1998). Resolution improvements in in vivo 1H NMR spectra with increased magnetic field strength. *Journal of magnetic resonance (San Diego, Calif. : 1997)*, *135*(1), 260–264. https://doi.org/10.1006/jmre.1998.1542

63. Hamilton, G., Middleton, M. S., Bydder, M., Yokoo, T., Schwimmer, J. B., Kono, Y., Patton, H. M., Lavine, J. E., & Sirlin, C. B. (2009). Effect of PRESS and STEAM sequences on magnetic resonance spectroscopic liver fat quantification. *Journal of Magnetic Resonance Imaging, 30*(1), 145–152. https://doi.org/10.1002/jmri.21809

64. Hancock, R., Pugh, K. R., & Hoeft, F. (2017). Neural Noise Hypothesis of Developmental Dyslexia. *Trends in cognitive sciences*, *21*(6), 434–448. https://doi.org/10.1016/j.tics.2017.03.008

65. Hand, L. J., Paterson, L. M., & Lingford-Hughes, A. R. (2024). Re-evaluating our focus in addiction: emotional dysregulation is a critical driver of relapse to drug use. *Translational psychiatry*, *14*(1), 467. https://doi.org/10.1038/s41398-024-03159-5

66. Harris, R. K., Becker, E. D., Cabral de Menezes, S. M., Goodfellow, R., & Granger, P. (2002). NMR nomenclature: Nuclear spin properties and conventions for chemical shifts: IUPAC recommendations 2001. *Solid State Nuclear Magnetic Resonance, 22*(4), 458–483. https://doi.org/10.1006/snmr.2002.0063

67. Heim, S., Alter, K., Ischebeck, A. K., Amunts, K., Eickhoff, S. B., Mohlberg, H., Zilles, K., von Cramon, D. Y., & Friederici, A. D. (2005). The role of the left Brodmann's areas 44 and 45 in reading words and pseudowords. *Brain research. Cognitive brain research*, *25*(3), 982–993. https://doi.org/10.1016/j.cogbrainres.2005.09.022

68. Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J. L., McMillon, G., Kolchugina, G., Black, J. M., Faizi, A., Deutsch, G. K., Siok, W. T., Reiss, A. L., Whitfield-Gabrieli, S., & Gabrieli, J. D. (2007). Functional and morphometric brain dissociation between dyslexia and reading ability. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(10), 4234–4239. https://doi.org/10.1073/pnas.0609399104

69. Hofmann, L., Slotboom, J., Jung, B., Maloca, P., Boesch, C., & Kreis, R. (2002). Quantitative ^{1}H-magnetic resonance spectroscopy of human brain: Influence of composition and parameterization of the basis set in linear combination model-fitting. *Magnetic Resonance in Medicine, 48*(3), 440–453. https://doi.org/10.1002/mrm.10246

70. Horowitz-Kraus, T., Brunst, K. J., & Cecil, K. M. (2018). Children With Dyslexia and Typical Readers: Sex-Based Choline Differences Revealed Using Proton Magnetic Resonance Spectroscopy Acquired Within Anterior Cingulate Cortex. *Frontiers in human neuroscience*, *12*, 466. https://doi.org/10.3389/fnhum.2018.00466

71. Huang, Z., Davis, H. H., IV, Yue, Q., Wiebking, C., Duncan, N. W., Zhang, J., Wagner, N. F., Wolff, A., & Northoff, G. (2015). Increase in glutamate/glutamine concentration in the medial prefrontal cortex during mental imagery: A combined functional mrs and fMRI study. *Human brain mapping*, *36*(8), 3204–3212. https://doi.org/10.1002/hbm.22841

72. Hui, S. C. N., Saleh, M. G., Zöllner, H. J., Oeltzschner, G., Fan, H., Li, Y., Song, Y., Jiang, H., Near, J., Lu, H., Mori, S., & Edden, R. A. E. (2022). MRSCloud: A cloud-based MRS tool for basis set simulation. *Magnetic resonance in medicine*, *88*(5), 1994–2004. https://doi.org/10.1002/mrm.29370

73. Ip, I. B., Berrington, A., Hess, A. T., Parker, A. J., Emir, U. E., & Bridge, H. (2017). Combined fMRI-MRS acquires simultaneous glutamate and BOLD-fMRI signals in the human brain. *NeuroImage*, *155*, 113–119. https://doi.org/10.1016/j.neuroimage.2017.04.030

74. Ip, I. B., Emir, U. E., Parker, A. J., Campbell, J., & Bridge, H. (2019). Comparison of Neurochemical and BOLD Signal Contrast Response Functions in the Human Visual Cortex. *The Journal of Neuroscience*, *39*(40), 7968–7975. https://doi.org/10.1523/JNEUROSCI.3021-18.2019

75. Jednoróg, K., Gawron, N., Marchewka, A., Heim, S., & Grabowska, A. (2014). Cognitive subtypes of dyslexia are characterized by distinct patterns of grey matter volume. *Brain structure & function*, *219*(5), 1697–1707. https://doi.org/10.1007/s00429-013-0595-6

76. Jelen, L. A., Lythgoe, D. J., Jackson, J. B., Howard, M. A., Stone, J. M., & Egerton, A. (2021). Imaging Brain Glx Dynamics in Response to Pressure Pain Stimulation: A ^{1}H-fMRS Study. *Frontiers in psychiatry*, *12*, 681419. https://doi.org/10.3389/fpsyt.2021.681419

77. Juchem, C., Cudalbu, C., de Graaf, R. A., Gruetter, R., Henning, A., Hetherington, H. P., & Boer, V. O. (2021). B0 shimming for in vivo magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR in Biomedicine, 34*(5), e4350. https://doi.org/10.1002/nbm.4350

78. Just, N., Xin, L., Frenkel, H., & Gruetter, R. (2013). Characterization of sustained BOLD activation in the rat barrel cortex and neurochemical consequences. *NeuroImage*, *74*, 343–351. https://doi.org/10.1016/j.neuroimage.2013.02.042

79. Kelter R. (2020). Bayesian alternatives to null hypothesis significance testing in biomedical research: a non-technical introduction to Bayesian inference with JASP. *BMC medical research methodology*, *20*(1), 142. https://doi.org/10.1186/s12874-020-00980-6

80. Keysers, C., Gazzola, V., & Wagenmakers, E. J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature neuroscience*, *23*(7), 788–799. https://doi.org/10.1038/s41593-020-0660-4

81. Kiemes, A., Davies, C., Kempton, M. J., Lukow, P. B., Bennallick, C., Stone, J. M., & Modinos, G. (2021). GABA, Glutamate and Neural Activity: A Systematic Review With Meta-Analysis of Multimodal ^{1}H-MRS-fMRI Studies. *Frontiers in psychiatry*, *12*, 644315. https://doi.org/10.3389/fpsyt.2021.644315

82. Klose U. (2008). Measurement sequences for single voxel proton MR spectroscopy. *European journal of radiology*, *67*(2), 194–201. https://doi.org/10.1016/j.ejrad.2008.03.023

83. Kniffin, A. R., & Briand, L. A. (2024). Sex differences in glutamate transmission and plasticity in reward related regions. *Frontiers in behavioral neuroscience*, *18*, 1455478. https://doi.org/10.3389/fnbeh.2024.1455478

84. Knouse, M. C., McGrath, A. G., Deutschmann, A. U., Rich, M. T., Zallar, L. J., Rajadhyaksha, A. M., & Briand, L. A. (2022). Sex differences in the medial prefrontal cortical glutamate system. *Biology of sex differences*, *13*(1), 66. https://doi.org/10.1186/s13293-022-00468-6

85. Kolasinski, J., Hinson, E. L., Divanbeighi Zand, A. P., Rizov, A., Emir, U. E., & Stagg, C. J. (2019). The dynamics of cortical GABA in human motor learning. *The Journal of physiology*, *597*(1), 271–282. https://doi.org/10.1113/JP276626

86. Koolschijn, R. S., Clarke, W. T., Ip, I. B., Emir, U. E., & Barron, H. C. (2023). Event-related functional magnetic resonance spectroscopy. *NeuroImage, 276,* 120194. https://doi.org/10.1016/j.neuroimage.2023.120194

87. Kossowski, B., Chyl, K., Kacprzak, A., Bogorodzki, P., & Jednoróg, K. (2019). Dyslexia and age related effects in the neurometabolites concentration in the visual and temporo-parietal cortex. *Scientific reports*, *9*(1), 5096. https://doi.org/10.1038/s41598-019-41473-x

88. Koyama, M. S., Di Martino, A., Kelly, C., Jutagir, D. R., Sunshine, J., Schwartz, S. J., Castellanos, F. X., & Milham, M. P. (2013). Cortical signatures of dyslexia and remediation: an intrinsic functional connectivity approach. *PloS one*, *8*(2), e55454. https://doi.org/10.1371/journal.pone.0055454

89. Krafnick, A. J., & Evans, T. M. (2019). Neurobiological Sex Differences in Developmental Dyslexia. *Frontiers in psychology*, *9*, 2669. https://doi.org/10.3389/fpsyg.2018.02669

90. Kreis R. (2004). Issues of spectral quality in clinical 1H-magnetic resonance spectroscopy and a gallery of artifacts. *NMR in biomedicine*, *17*(6), 361–381. https://doi.org/10.1002/nbm.891

91. Kurcyus, K., Annac, E., Hanning, N. M., Harris, A. D., Oeltzschner, G., Edden, R., & Riedl, V. (2018). Opposite Dynamics of GABA and Glutamate Levels in the Occipital Cortex during Visual Processing. *The Journal of Neuroscience*, *38*(46), 9967–9976. https://doi.org/10.1523/JNEUROSCI.1214-18.2018

92. Landheer, K., & Juchem, C. (2021). FAMASITO: FASTMAP Shim Tool towards user-friendly single-step B_0 homogenization. *NMR in biomedicine*, *34*(6), e4486. https://doi.org/10.1002/nbm.4486

93. Landheer, K., Swanberg, K. M., & Juchem, C. (2021). Magnetic resonance Spectrum simulator (MARSS), a novel software package for fast and computationally efficient basis set simulation. *NMR in biomedicine*, *34*(5), e4129. https://doi.org/10.1002/nbm.4129

94. Laycock, S. K., Wilkinson, I. D., Wallis, L. I., Darwent, G., Wonders, S. H., Fawcett, A. J., Griffiths, P. D., & Nicolson, R. I. (2008). Cerebellar volume and cerebellar metabolic characteristics in adults with dyslexia. *Annals of the New York Academy of Sciences*, *1145*, 222–236. https://doi.org/10.1196/annals.1416.002

95. Lea-Carnall, C. A., El-Deredy, W., Stagg, C. J., Williams, S. R., & Trujillo-Barreto, N. J. (2023). A mean-field model of glutamate and GABA synaptic dynamics for functional MRS. *NeuroImage*, *266*, 119813. https://doi.org/10.1016/j.neuroimage.2022.119813

96. Lebel, C., MacMaster, F. P., & Dewey, D. (2016). Brain metabolite levels and language abilities in preschool children. *Brain and behavior*, *6*(10), e00547. https://doi.org/10.1002/brb3.547

97. Li, Y., Xu, D., Ozturk-Isik, E., Lupo, J. M., Chen, A. P., Vigneron, D. B., & Nelson, S. J. (2012). T1 and T2 metabolite relaxation times in normal brain at 3T and 7T. *Journal of Molecular Imaging & Dynamics, S1*, 002. https://doi.org/10.4172/2155-9937.S1-002

98. Lin, A., Andronesi, O., Bogner, W., Choi, I. Y., Coello, E., Cudalbu, C., Juchem, C., Kemp, G. J., Kreis, R., Krššák, M., Lee, P., Maudsley, A. A., Meyerspeer, M., Mlynarik, V., Near, J., Öz, G., Peek, A. L., Puts, N. A., Ratai, E. M., Tkáč, I., … Experts' Working Group on Reporting Standards for MR Spectroscopy (2021). Minimum Reporting Standards for in vivo Magnetic Resonance Spectroscopy (MRSinMRS): Experts' consensus recommendations. *NMR in biomedicine*, *34*(5), e4484. https://doi.org/10.1002/nbm.4484

99. Liu, Y., Chen, H., Zhang, Z., & Tal, A. (2025). 2D spectral-temporal fitting of functional MRS improves the fitting precision and noise robustness. *Biomedical Signal Processing and Control, 109,* 108018. https://doi.org/10.1016/j.bspc.2025.108018

100. Lynn, J., Woodcock, E. A., Anand, C., Khatib, D., & Stanley, J. A. (2018). Differences in steady-state glutamate levels and variability between 'non-task-active' conditions: Evidence from ^{1}H fMRS of the prefrontal cortex. *NeuroImage*, *172*, 554–561. https://doi.org/10.1016/j.neuroimage.2018.01.069

101. Magistretti, P. J., & Allaman, I. (2015). A cellular perspective on brain energy metabolism and functional imaging. *Neuron*, *86*(4), 883–901. https://doi.org/10.1016/j.neuron.2015.03.035

102. Mangia, S., Tkáč, I., Gruetter, R., Van De Moortele, P. F., Giove, F., Maraviglia, B., & Uğurbil, K. (2006). Sensitivity of single-voxel 1H-MRS in investigating the metabolism of the activated human visual cortex at 7 T. *Magnetic resonance imaging*, *24*(4), 343–348. https://doi.org/10.1016/j.mri.2005.12.023

103. Mangia, S., Tkáč, I., Gruetter, R., Van de Moortele, P. F., Maraviglia, B., & Uğurbil, K. (2007). Sustained neuronal activation raises oxidative metabolism to a new steady-state level: evidence from 1H NMR spectroscopy in the human visual cortex. *Journal of Cerebral Blood Flow & Metabolism*, *27*(5), 1055–1063. https://doi.org/10.1038/sj.jcbfm.9600401

104. Martin, A., Kronbichler, M., & Richlan, F. (2016). Dyslexic brain activation abnormalities in deep and shallow orthographies: A meta-analysis of 28 functional neuroimaging studies. *Human brain mapping*, *37*(7), 2676–2699. https://doi.org/10.1002/hbm.23202

105. Maudsley, A. A., Andronesi, O. C., Barker, P. B., Bizzi, A., Bogner, W., Henning, A., Nelson, S. J., Posse, S., Shungu, D. C., & Soher, B. J. (2021). Advanced magnetic resonance spectroscopic neuroimaging: Experts' consensus recommendations. *NMR in biomedicine*, *34*(5), e4309. https://doi.org/10.1002/nbm.4309

106. Maudsley, A. A., Andronesi, O. C., Barker, P. B., Bizzi, A., Bogner, W., Henning, A., Nelson, S. J., Posse, S., Shungu, D. C., & Soher, B. J. (2021). Advanced magnetic resonance spectroscopic neuroimaging: Experts' consensus recommendations. *NMR in Biomedicine, 34*(5), e4309. https://doi.org/10.1002/nbm.4309

107. McLean, M. A., Woermann, F. G., Barker, G. J., & Duncan, J. S. (2000). Quantitative analysis of short echo time (1)H-MRSI of cerebral gray and white matter. *Magnetic resonance in medicine*, *44*(3), 401–411. https://doi.org/10.1002/1522-2594(200009)44:3<401::aid-mrm10>3.0.co;2-w

108. Mekle, R., Mlynárik, V., Gambarota, G., Hergt, M., Krueger, G., & Gruetter, R. (2009). MR spectroscopy of the human brain with enhanced signal intensity at ultrashort echo times on a clinical platform at 3T and 7T. *Magnetic resonance in medicine*, *61*(6), 1279–1285. https://doi.org/10.1002/mrm.21961

109. Mescher, M., Merkle, H., Kirsch, J., Garwood, M., & Gruetter, R. (1998). Simultaneous in vivo spectral editing and water suppression. *NMR in biomedicine*, *11*(6), 266–272. https://doi.org/10.1002/(sici)1099-1492(199810)11:6<266::aid-nbm530>3.0.co;2-j

110. Mocioiu, V., Ortega-Martorell, S., Olier, I., Jablonski, M., Starcukova, J., Lisboa, P., Arús, C., & Julià-Sapé, M. (2015). From raw data to data-analysis for magnetic resonance spectroscopy – the missing link: jMRUI2XML. *BMC Bioinformatics, 16*, 378. https://doi.org/10.1186/s12859-015-0796-5
111. Moonen, C. T., von Kienlin, M., van Zijl, P. C., Cohen, J., Gillen, J., Daly, P., & Wolf, G. (1989). Comparison of single-shot localization methods (STEAM and PRESS) for in vivo proton NMR spectroscopy. *NMR in biomedicine*, *2*(5-6), 201–208. https://doi.org/10.1002/nbm.1940020506
112. Mullins P. G. (2018). Towards a theory of functional magnetic resonance spectroscopy (fMRS): A meta-analysis and discussion of using MRS to measure changes in neurotransmitters in real time. *Scandinavian journal of psychology*, *59*(1), 91–103. https://doi.org/10.1111/sjop.12411
113. Nakai, T., & Okanoya, K. (2016). Individual variability in verbal fluency correlates with γ-aminobutyric acid concentration in the left inferior frontal gyrus. *Neuroreport*, *27*(13), 987–991. https://doi.org/10.1097/WNR.0000000000000645
114. Narayana, P. A., Fotedar, L. K., Jackson, E. F., Bohan, T. P., Butler, I. J., & Wolinsky, J. S. (1989). Regional in vivo proton magnetic resonance spectroscopy of brain. *Journal of Magnetic Resonance, 83*(1), 44–52. https://doi.org/10.1016/0022-2364(89)90290-4
115. Near, J., Harris, A. D., Juchem, C., Kreis, R., Marjańska, M., Öz, G., Slotboom, J., Wilson, M., & Gasparovic, C. (2021). Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR in Biomedicine, 34*(5), e4257. https://doi.org/10.1002/nbm.4257
116. Oeltzschner, G., Zöllner, H. J., Hui, S. C. N., Mikkelsen, M., Saleh, M. G., Tapper, S., & Edden, R. A. E. (2020). Osprey: Open-source processing, reconstruction & estimation of magnetic resonance spectroscopy data. *Journal of neuroscience methods*, *343*, 108827. https://doi.org/10.1016/j.jneumeth.2020.108827
117. Öz, G., Deelchand, D. K., Wijnen, J. P., Mlynárik, V., Xin, L., Mekle, R., Noeske, R., Scheenen, T. W. J., Tkáč, I., & Experts' Working Group on Advanced Single Voxel 1H MRS. (2021). Advanced single voxel ^{1}H magnetic resonance spectroscopy techniques in humans: Experts' consensus recommendations. *NMR in Biomedicine, 34*(5), e4236. https://doi.org/10.1002/nbm.4236

118. Pasanta, D., He, J. L., Ford, T., Oeltzschner, G., Lythgoe, D. J., & Puts, N. A. (2023). Functional MRS studies of GABA and glutamate/Glx - A systematic review and meta-analysis. *Neuroscience and biobehavioral reviews*, *144*, 104940. https://doi.org/10.1016/j.neubiorev.2022.104940

119. Pasanta, D., Htun, K. T., Pan, J., Tungjai, M., Kaewjaeng, S., Kim, H., Kaewkhao, J., & Kothan, S. (2021). Magnetic Resonance Spectroscopy of Hepatic Fat from Fundamental to Clinical Applications. Diagnostics, 11(5), 842. https://doi.org/10.3390/diagnostics11050842

120. Pasanta, D., White, D. J., He, J. L., Ford, T. C., & Puts, N. A. (2024). GABA and glutamate response to social processing: a functional MRS feasibility study. *NMR in biomedicine*, *37*(7), e5092. https://doi.org/10.1002/nbm.5092

121. Peek, A. L., Rebbeck, T. J., Leaver, A. M., Foster, S. L., Refshauge, K. M., Puts, N. A., Oeltzschner, G., & MRS Expert Panel. (2023). A comprehensive guide to MEGA-PRESS for GABA measurement. *Analytical Biochemistry, 669*, 115113. https://doi.org/10.1016/j.ab.2023.115113

122. Peek, A. L., Rebbeck, T., Puts, N. A., Watson, J., Aguila, M. R., & Leaver, A. M. (2020). Brain GABA and glutamate levels across pain conditions: A systematic literature review and meta-analysis of 1H-MRS studies using the MRS-Q quality assessment tool. *NeuroImage*, *210*, 116532. https://doi.org/10.1016/j.neuroimage.2020.116532

123. Perica, M. I., Calabro, F. J., Larsen, B., Foran, W., Yushmanov, V. E., Hetherington, H., Tervo-Clemmens, B., Moon, C. H., & Luna, B. (2022). Development of frontal GABA and glutamate supports excitation/inhibition balance from adolescence into adulthood. *Progress in neurobiology*, *219*, 102370. https://doi.org/10.1016/j.pneurobio.2022.102370

124. Posse, S., Otazo, R., Dager, S. R., & Alger, J. (2013). MR spectroscopic imaging: Principles and recent advances. *Journal of Magnetic Resonance Imaging, 37*(6), 1301–1325. https://doi.org/10.1002/jmri.23945

125. Pradhan, S., Bonekamp, S., Gillen, J. S., Rowland, L. M., Wijtenburg, S. A., Edden, R. A. E., & Barker, P. B. (2015). Comparison of single voxel brain MRS at 3T and 7T using 32-channel head coils. *Magnetic Resonance Imaging, 33*(9), 1013–1018. https://doi.org/10.1016/j.mri.2015.06.003

126. Prichard, J., Rothman, D., Novotny, E., Petroff, O., Kuwabara, T., Avison, M., Howseman, A., Hanstock, C., & Shulman, R. (1991). Lactate rise detected by 1H NMR in human visual cortex during physiologic stimulation. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(13), 5829–5831. https://doi.org/10.1073/pnas.88.13.5829

127. Provencher, S. W. (1993). Estimation of metabolite concentrations from localized in vivo proton NMR spectra. *Magnetic Resonance in Medicine, 30*(6), 672–679. https://doi.org/10.1002/mrm.1910300604

128. Provencher, S. W. (2001). Automatic quantitation of localized in vivo ^{1}H spectra with LCModel. *NMR in Biomedicine, 14*(4), 260–264. https://doi.org/10.1002/nbm.698

129. Pugh, K. R., Frost, S. J., Rothman, D. L., Hoeft, F., Del Tufo, S. N., Mason, G. F., Molfese, P. J., Mencl, W. E., Grigorenko, E. L., Landi, N., Preston, J. L., Jacobsen, L., Seidenberg, M. S., & Fulbright, R. K. (2014). Glutamate and choline levels predict individual differences in reading ability in emergent readers. *The Journal of Neuroscience*, *34*(11), 4082–4089. https://doi.org/10.1523/JNEUROSCI.3907-13.2014

130. Puts, N. A., & Edden, R. A. (2012). In vivo magnetic resonance spectroscopy of GABA: a methodological review. *Progress in nuclear magnetic resonance spectroscopy*, *60*, 29–41. https://doi.org/10.1016/j.pnmrs.2011.06.001

131. Rae C. D. (2014). A guide to the metabolic pathways and function of metabolites observed in human brain 1H magnetic resonance spectra. *Neurochemical research*, *39*(1), 1–36. https://doi.org/10.1007/s11064-013-1199-5

132. Rae, C., Lee, M. A., Dixon, R. M., Blamire, A. M., Thompson, C. H., Styles, P., Talcott, J., Richardson, A. J., & Stein, J. F. (1998). Metabolic abnormalities in developmental dyslexia detected by 1H magnetic resonance spectroscopy. *Lancet (London, England)*, *351*(9119), 1849–1852. https://doi.org/10.1016/S0140-6736(97)99001-2

133. Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., & Scotto di Covella, L. (2018). Neuroanatomy of developmental dyslexia: Pitfalls and promise. *Neuroscience and biobehavioral reviews*, *84*, 434–452. https://doi.org/10.1016/j.neubiorev.2017.08.001

134. Ramus, F., Marshall, C. R., Rosen, S., & van der Lely, H. K. (2013). Phonological deficits in specific language impairment and developmental dyslexia: towards a multidimensional model. *Brain : a journal of neurology*, *136*(Pt 2), 630–645. https://doi.org/10.1093/brain/aws356

135. Richards, T. L., Dager, S. R., Corina, D., Serafini, S., Heide, A. C., Steury, K., Strauss, W., Hayes, C. E., Abbott, R. D., Craft, S., Shaw, D., Posse, S., & Berninger, V. W. (1999). Dyslexic children have abnormal brain lactate response to reading-related language tasks. *AJNR. American journal of neuroradiology*, *20*(8), 1393–1398.

136. Richardson, A. J., Cox, I. J., Sargentoni, J., & Puri, B. K. (1997). Abnormal cerebral phospholipid metabolism in dyslexia indicated by phosphorus-31 magnetic resonance spectroscopy. *NMR in biomedicine*, *10*(7), 309–314. https://doi.org/10.1002/(sici)1099-1492(199710)10:7<309::aid-nbm484>3.0.co;2-0

137. Richlan F. (2019). The Functional Neuroanatomy of Letter-Speech Sound Integration and Its Relation to Brain Abnormalities in Developmental Dyslexia. *Frontiers in human neuroscience*, *13*, 21. https://doi.org/10.3389/fnhum.2019.00021

138. Richlan F. (2020). The Functional Neuroanatomy of Developmental Dyslexia Across Languages and Writing Systems. *Frontiers in psychology*, *11*, 155. https://doi.org/10.3389/fpsyg.2020.00155

139. Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: a quantitative meta-analysis of neuroimaging studies. *Human brain mapping*, *30*(10), 3299–3308. https://doi.org/10.1002/hbm.20752

140. Roid, G.H., Sajewicz-Radtke, U., Radtke, B.M., & Lipowska, M. (2017). Skala Inteligencji Stanford-Binet, Edycja Piąta [Stanford-Binet Intelligence Scales, Fifth Edition]. Pracownia Testów Psychologicznych i Pedagogicznych [Laboratory of Psychological and Pedagogical Tests]: Gdańsk, Poland.

141. Ross, B. D., Lin, A., Harris, K., Bhattacharya, P., & Schweinsburg, B. (2003). Clinical experience with ^{13}C MRS in vivo. NMR in Biomedicine, 16(6–7), 358–369. https://doi.org/10.1002/nbm.852

142. Rothman, D. L., Behar, K. L., Hyder, F., & Shulman, R. G. (2003). In vivo NMR studies of the glutamate neurotransmitter flux and neuroenergetics: implications for brain function. *Annual review of physiology*, *65*, 401–427. https://doi.org/10.1146/annurev.physiol.65.092101.142131
143. Sacolick, L. I., Rothman, D. L., & de Graaf, R. A. (2007). Adiabatic refocusing pulses for volume selection in magnetic resonance spectroscopic imaging. *Magnetic Resonance in Medicine, 57*(4), 548–553. https://doi.org/10.1002/mrm.21162
144. Santos-Díaz, A., & Noseworthy, M. D. (2020). Phosphorus magnetic resonance spectroscopy and imaging (31P-MRS/MRSI) as a window to brain and muscle metabolism: A review of the methods. *Biomedical Signal Processing and Control, 60,* 101967. https://doi.org/10.1016/j.bspc.2020.101967
145. Schaller, B., Xin, L., O'Brien, K., Magill, A. W., & Gruetter, R. (2014). Are glutamate and lactate increases ubiquitous to physiological activation? A (1)H functional MR spectroscopy study during motor activation in human brain at 7Tesla. *NeuroImage*, *93 Pt 1*, 138–145. https://doi.org/10.1016/j.neuroimage.2014.02.016
146. Schwarz, J., Lizarazu, M., Lallier, M., & Klimovich-Gray, A. (2024). Phonological deficits in dyslexia impede lexical processing of spoken words: Linking behavioural and MEG data. *Cortex; a journal devoted to the study of the nervous system and behavior*, *171*, 204–222. https://doi.org/10.1016/j.cortex.2023.10.003
147. Shokry, A. (2012). MRS of brain tumors: Diagrammatic representations and diagnostic approach. *The Egyptian Journal of Radiology and Nuclear Medicine, 43*(4), 603–612. https://doi.org/10.1016/j.ejrnm.2012.07.006
148. Sibson, N. R., Dhankhar, A., Mason, G. F., Rothman, D. L., Behar, K. L., & Shulman, R. G. (1998). Stoichiometric coupling of brain glucose metabolism and glutamatergic neuronal activity. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(1), 316–321. https://doi.org/10.1073/pnas.95.1.316
149. Simpson, R., Devenyi, G. A., Jezzard, P., Hennessy, T. J., & Near, J. (2017). Advanced processing and simulation of MRS data using the FID appliance (FID-A)-An open source, MATLAB-based toolkit. *Magnetic resonance in medicine*, *77*(1), 23–33. https://doi.org/10.1002/mrm.26091

150. Skoch, A., Jiru, F., & Bunke, J. (2008). Spectroscopic imaging: basic principles. *European journal of radiology*, *67*(2), 230–239. https://doi.org/10.1016/j.ejrad.2008.03.003
151. Snowling M. J. (2001). From language to reading and dyslexia. *Dyslexia (Chichester, England)*, *7*(1), 37–46. https://doi.org/10.1002/dys.185
152. Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological bulletin*, *142*(5), 498–545. https://doi.org/10.1037/bul0000037
153. Soher, B. J., Semanchuk, P., Todd, D., Ji, X., Deelchand, D., Joers, J., Oz, G., & Young, K. (2023). Vespa: Integrated applications for RF pulse design, spectral simulation and MRS data analysis. *Magnetic resonance in medicine*, *90*(3), 823–838. https://doi.org/10.1002/mrm.29686
154. Stamatelatou, A., Scheenen, T. W. J., & Heerschap, A. (2022). Developments in proton MR spectroscopic imaging of prostate cancer. *Magnetic Resonance Materials in Physics, Biology and Medicine, 35*(4), 645–665. https://doi.org/10.1007/s10334-022-01011-9
155. Stanley, J. A., & Raz, N. (2018). Functional magnetic resonance spectroscopy: The "new" MRS for cognitive neuroscience and psychiatry research. *Frontiers in Psychiatry, 9*, 76. https://doi.org/10.3389/fpsyt.2018.00076
156. Stefan, D., Cesare, F. D., Andrasescu, A., Popa, E., Lazariev, A., Vescovo, E., Strbak, O., Williams, S., Starcuk, Z., Cabanas, M., van Ormondt, D., & Graveron-Demilly, D. (2009). Quantitation of magnetic resonance spectroscopy signals: The jMRUI software package. *Measurement Science and Technology, 20*(10), 104035. https://doi.org/10.1088/0957-0233/20/10/104035
157. Szczerbiński, M., & Pelc-Pękała, O. (2013). Zestaw metod do diagnozy trudności w czytaniu [A Set of Tools for Diagnosing Reading Difficulties]. Pracownia Testów Psychologicznych i Pedagogicznych [Laboratory of Psychological and Pedagogical Tests]: Gdańsk, Poland.
158. Takado, Y., Takuwa, H., Sampei, K., Urushihata, T., Takahashi, M., Shimojo, M., Uchida, S., Nitta, N., Shibata, S., Nagashima, K., Ochi, Y., Ono, M., Maeda, J., Tomita, Y., Sahara, N., Near, J., Aoki, I., Shibata, K., & Higuchi, M. (2022). MRS-measured glutamate versus GABA reflects excitatory versus inhibitory neural activities in awake mice. *Journal of Cerebral Blood Flow & Metabolism*, *42*(1), 197–212. https://doi.org/10.1177/0271678X211045449

159. Tannús, A., & Garwood, M. (1997). Adiabatic pulses. *NMR in Biomedicine, 10*(8), 423–434. https://doi.org/10.1002/(SICI)1099-1492(199712)10:8<423::AID-NBM488>3.0.CO;2-X

160. Terpstra, M., Cheong, I., Lyu, T., Deelchand, D. K., Emir, U. E., Bednařík, P., Eberly, L. E., & Öz, G. (2016). Test-retest reproducibility of neurochemical profiles with short-echo, single-voxel MR spectroscopy at 3T and 7T. *Magnetic resonance in medicine*, *76*(4), 1083–1091. https://doi.org/10.1002/mrm.26022

161. Tkáč, I., Deelchand, D., Dreher, W., Hetherington, H., Kreis, R., Kumaragamage, C., Považan, M., Spielman, D. M., Strasser, B., & de Graaf, R. A. (2021). Water and lipid suppression techniques for advanced 1H MRS and MRSI of the human brain: Experts' consensus recommendations. *NMR in Biomedicine, 34*(5), e4459. https://doi.org/10.1002/nbm.4459

162. Tkáč, I., Oz, G., Adriany, G., Uğurbil, K., & Gruetter, R. (2009). In vivo 1H NMR spectroscopy of the human brain at high magnetic fields: metabolite quantification at 4T vs. 7T. *Magnetic resonance in medicine*, *62*(4), 868–879. https://doi.org/10.1002/mrm.22086

163. Tognarelli, J. M., Dawood, M., Shariff, M. I. F., Grover, V. P. B., Crossey, M. M. E., Cox, I. J., Taylor-Robinson, S. D., & McPhail, M. J. W. (2015). Magnetic resonance spectroscopy: Principles and techniques: Lessons for clinicians. *Journal of Clinical and Experimental Hepatology, 5*(4), 320–328. https://doi.org/10.1016/j.jceh.2015.10.006

164. Turker, S., Fumagalli, B., Kuhnke, P., & Hartwigsen, G. (2025). The 'reading' brain: Meta-analytic insight into functional activation during reading in adults. *Neuroscience and biobehavioral reviews*, *173*, 106166. https://doi.org/10.1016/j.neubiorev.2025.106166

165. Turri, C., Di Dona, G., Santoni, A., Zamfira, D. A., Franchin, L., Melcher, D., & Ronconi, L. (2023). Periodic and Aperiodic EEG Features as Potential Markers of Developmental Dyslexia. *Biomedicines*, *11*(6), 1607. https://doi.org/10.3390/biomedicines11061607

166. Valtcheva, S., & Venance, L. (2019). Control of Long-Term Plasticity by Glutamate Transporters. *Frontiers in synaptic neuroscience*, *11*, 10. https://doi.org/10.3389/fnsyn.2019.00010

167. van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, *43*(2), 271–282. https://doi.org/10.1016/j.neuron.2004.06.025

168. Vidal, C., Content, A., & Chetail, F. (2017). BACS: The Brussels Artificial Character Sets for studies in cognitive psychology and neuroscience. *Behavior research methods*, *49*(6), 2093–2112. https://doi.org/10.3758/s13428-016-0844-8

169. Vidyasagar, T. R., & Pammer, K. (2010). Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in cognitive sciences*, *14*(2), 57–63. https://doi.org/10.1016/j.tics.2009.12.003

170. Weinberg, B. D., Kuruva, M., Shim, H., & Mullins, M. E. (2021). Clinical applications of magnetic resonance spectroscopy in brain tumors: From diagnosis to treatment. *Radiologic Clinics of North America, 59*(3), 349–362. https://doi.org/10.1016/j.rcl.2021.01.004

171. Wilmot, A., Hasking, P., Leitão, S., Hill, E., & Boyes, M. (2023). Understanding Mental Health in Developmental Dyslexia: A Scoping Review. *International journal of environmental research and public health*, *20*(2), 1653. https://doi.org/10.3390/ijerph20021653

172. Wilson, M., Andronesi, O., Barker, P. B., Bartha, R., Bizzi, A., Bolan, P. J., Brindle, K. M., Choi, I. Y., Cudalbu, C., Dydak, U., Emir, U. E., Gonzalez, R. G., Gruber, S., Gruetter, R., Gupta, R. K., Heerschap, A., Henning, A., Hetherington, H. P., Huppi, P. S., Hurd, R. E., … Howe, F. A. (2019). Methodological consensus on clinical proton MRS of the brain: Review and recommendations. *Magnetic resonance in medicine*, *82*(2), 527–550. https://doi.org/10.1002/mrm.27742

173. Wilson, M., Reynolds, G., Kauppinen, R. A., Arvanitis, T. N., & Peet, A. C. (2011). A constrained least-squares approach to the automated quantitation of in vivo ^{1}H magnetic resonance spectroscopy data. *Magnetic Resonance in Medicine, 65*(1), 1–12. https://doi.org/10.1002/mrm.22579

174. Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*(3), 415–438. https://doi.org/10.1037/0022-0663.91.3.415

175. Woodcock, E. A., Anand, C., Khatib, D., Diwadkar, V. A., & Stanley, J. A. (2018). Working Memory Modulates Glutamate Levels in the Dorsolateral Prefrontal Cortex during ^{1}H fMRS. *Frontiers in psychiatry*, *9*, 66. https://doi.org/10.3389/fpsyt.2018.00066

176. Wyss, P. O., Hock, A., & Kollias, S. (2017). The application of human spinal cord magnetic resonance spectroscopy to clinical studies: A review. *Seminars in Ultrasound, CT and MRI, 38*(3), 153–162. https://doi.org/10.1053/j.sult.2016.07.005

177. Xu, P., Chen, A., Li, Y., Xing, X., & Lu, H. (2019). Medial prefrontal cortex in neurological diseases. *Physiological genomics*, *51*(9), 432–442. https://doi.org/10.1152/physiolgenomics.00006.2019

178. Xu, S., Pratt, S. J. P., Spangenburg, E. E., & Lovering, R. M. (2012). Early metabolic changes measured by 1H MRS in healthy and dystrophic muscle after injury. *Journal of Applied Physiology, 113*(5), 808–816. https://doi.org/10.1152/japplphysiol.00530.2012

179. Yakovlev, A., Manzhurtsev, A., Menshchikov, P., Ublinskiy, M., Melnikov, I., Kupriyanov, D., Akhadov, T., & Semenova, N. (2022). Functional magnetic resonance spectroscopy study of total glutamate and glutamine in the human visual cortex activated by a short stimulus. *Biophysics, 67*(2), 265–273. https://doi.org/10.1134/S0006350922020245

180. Yan, X., Jiang, K., Li, H., Wang, Z., Perkins, K., & Cao, F. (2021). Convergent and divergent brain structural and functional abnormalities associated with developmental dyslexia. *eLife*, *10*, e69523. https://doi.org/10.7554/eLife.69523

181. Yang, L., Li, C., Li, X., Zhai, M., An, Q., Zhang, Y., Zhao, J., & Weng, X. (2022). Prevalence of Developmental Dyslexia in Primary School Children: A Systematic Review and Meta-Analysis. *Brain sciences*, *12*(2), 240. https://doi.org/10.3390/brainsci12020240

182. Ye, Z., Rüsseler, J., Gerth, I., & Münte, T. F. (2017). Audiovisual speech integration in the superior temporal region is dysfunctional in dyslexia. *Neuroscience*, *356*, 1–10. https://doi.org/10.1016/j.neuroscience.2017.05.017

183. Zhao, J., Yang, H., Weng, X., & Wang, Z. (2018). Emergent Attentional Bias Toward Visual Word Forms in the Environment: Evidence From Eye Movements. *Frontiers in psychology*, 9, 1378. https://doi.org/10.3389/fpsyg.2018.01378

List of Publications

During my doctoral studies, I co-authored the following publications:

1. Glica, A., Wasilewska, K., Kossowski, B., Żygierewicz, J., & Jednoróg, K. (2024). Sex Differences in Low-Level Multisensory Integration in Developmental Dyslexia. *The Journal of Neuroscience*, *44*(3), e0944232023. https://doi.org/10.1523/JNEUROSCI.0944-23.2023
2. Glica, A., Wasilewska, K., Jurkowska, J., Żygierewicz, J., Kossowski, B., & Jednoróg, K. (2025). Reevaluating the neural noise in dyslexia using biomarkers from electroencephalography and high-resolution magnetic resonance spectroscopy. *eLife, 13,* RP99920. https://doi.org/10.7554/eLife.99920

Conference Presentations

I also shared my research at the following conferences:

Oral presentations

1. "Does visual stimulation always lead to an increase in glutamate concentration?" at the MRS Workshop (Switzerland, 2022)
2. "Investigating Neural Noise in Dyslexia Using Functional Magnetic Resonance Spectroscopy (fMRS) and EEG Power Spectrum" at the NEURONUS Conference (Poland, 2022)
3. "Functional Magnetic Resonance Spectroscopy (fMRS) Study on Reading-Related Metabolite Changes at 7T" at the Symposium on fMRS in Bangor (Great Britain, 2025)

Poster presentations

4. "fMRS Study on Reading-Related Metabolite Changes" at the NEURONUS Conference (Poland, 2024)
5. "Functional Magnetic Resonance Spectroscopy (fMRS) Study on Reading-Related Metabolite Changes" at the Society for the Neurobiology of Language (SNL) Conference (Australia, 2024)