Characterizing the Extended Language Network in Individuals with Multiple Sclerosis

Author(s): Alexander S. Ratzan[1], Leila Simani[1], Jordan D. Dworkin[2], Korhan Buyukturkoglu[3], Claire S. Riley[3], Victoria M. Leavitt[1]

Corresponding Author:
Victoria M. Leavitt, vl2337@cumc.columbia.edu

Affiliation Information for all Authors: 1. Cognitive Neuroscience Division, Department of Neurology, Columbia University Irving Medical Center, New York, NY, 2. Department of Psychiatry, Columbia University and the New York State Psychiatric Institute, New York, NY, 3. The Center for Translational and Computational Neuroimmunology, Columbia University MS Center, New York, NY

Equal Author Contribution:

Contributions:
Alexander S. Ratzan: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Analysis or interpretation of data.
Leila Simani: Drafting/revision of the manuscript for content, including medical writing for content.
Jordan D. Dworkin: Drafting/revision of the manuscript for content, including medical writing for content; Analysis or interpretation of data.
Korhan Buyukturkoglu: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data.
Claire S. Riley: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data.
Victoria M. Leavitt: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data.

Figure Count:
4

Table Count:
3

Acknowledgment: The authors thank Lauren Heuer for assistance in preparing the manuscript for publication, and thank the individuals who participated in the study.

Study Funding: United States Department of Defense Congressionally Directed Medical Research Program (W81XWH-20-1-0503).

Disclosure:
VML has been compensated for advisory or consulting services by the following entities in the last year:

Novartis, Biogen. CSR has been compensated for advisory or consulting services by the following entities in the last year: EMD Serono, TG Therapeutics, Horizon, Novartis, Viracta, Genentech. ASR has nothing to disclose. JDD has nothing to disclose. KB has nothing to disclose.LS has nothing to disclose.

Abstract:

Background: Cognitive impairment is a pervasive, functionally limiting symptom of multiple sclerosis (MS), a disease of the central nervous system that is the most common non-traumatic cause of neurologic disability in young adults. Recently, language dysfunction has received increased attention as a prevalent and early affected cognitive domain in individuals with MS.
Objectives: To establish a network-level model of language dysfunction in MS.
Methods: Cognitive data and 3T structural and functional brain magnetic resonance imaging (MRI) scans were acquired from 54 MS patients and 54 healthy controls (HCs). Summary measures of the extended language network (ELN) and structural imaging metrics were calculated. Group differences in ELN summary measures were evaluated. Associations between ELN summary measures and language performance were assessed in both groups; in the MS group, a two-step regression analysis was applied to assess relationships between additional language-specific imaging measures and language performance.
Results: In comparison to the HC group, the MS group performed significantly worse on the semantic fluency and rapid automized naming tests ($p < 0.005$). Concerning the ELN summary measures, the MS group exhibited higher within-ELN connectivity than the HCs (0.11 ± 0.02 vs. 0.10 ± 0.01, $p < 0.05$, respectively). While no significant relationships between ELN summary measures and language function were observed in either group, the regression analysis identified a set of 17 imaging features that predicted performance on the rapid automized naming test ($p < 0.05$) and identified key white matter tracts predicting language function in individuals with MS.
Conclusion: The derived functional network-level measures, combined with the identified structural neuroimaging metrics, constitute a comprehensive set of imaging features to characterize language dysfunction in MS. Further studies leveraging these features may uncover underlying mechanisms and clinically relevant predictors of language dysfunction, potentially leading to improved precision treatment strategies for cognitively impaired patients with multiple sclerosis.

Keyword: Network model, Language dysfunction, Extended language network, Multiple sclerosis

Introduction

Multiple sclerosis (MS) is a chronic neurological disease characterized by neurodegeneration and axonal demyelination [1]. Cognitive impairment affects many individuals with MS and can arise early in the disease course [2]. Memory decline and slowed information processing speed are generally considered to be the most prominent features of cognitive impairment in MS [3-5]. Language dysfunction has only recently begun to receive widespread attention as an important affected domain [6-9]. In their seminal study by Rao et al.1991, performance on a verbal fluency test was identified as one of the most impaired cognitive measures in people with MS (pwMS). However, this test was characterized as a measure of recent memory, potentially leading to an important oversight in the field's conceptualization of the prominence of language dysfunction in MS [10]. Recently, a test of rapid automatized naming was the only objective cognitive test measure (of 9 different measures) that distinguished recently diagnosed pwMS from matched healthy controls (HCs), and word finding difficulties were the most commonly reported cognitive issue by pwMS early in their disease course [11]. This growing recognition of the prominence of language deficits in the cognitive profile of pwMS highlights the need for a mechanistic model to elucidate the cause(s) of disrupted language function in MS. The focus of this study is to provide an initial network-level model of language dysfunction for MS.

Traditionally, studies investigating the neural substrates of language (in the non-MS literature) have focused on mapping individual cortical regions to specific language functions (i.e., one-to-one brain-behavior relationships). Towards a network-level conceptualization of language function, Tomasi & Volkow employed resting-state functional magnetic resonance imaging (fMRI) to identify the extended language network (ELN) in 970 healthy adults [12]. The ELN is highly reproducible both during resting-state and task-based fMRI, recommending its use as a promising network model to explore language dysfunction in normal aging and clinical populations (e.g., temporal lobe epilepsy) [13-16]. In the context of MS, the few studies to date that have evaluated the neural substrates of language function examined relationships to cortical thickness and white matter microstructure [11,17]. Developing a network-level model of language for MS will permit mechanistic insights into this key and functionally limiting cognitive deficit.

Here, we utilized the ELN as a framework to develop a network-level model of language impairment in MS. Applying an established approach for characterizing network (re)organization of functionally specific subnetworks using resting state functional connectivity (rsFC), we derived language-specific rsFC summary measures to capture non-random patterns of network-level reorganization of the language network: within-ELN connectivity, between-ELN connectivity, segregation index (Seg-I), and anteriority index (Ant-I) [18,19]. We then tested: (a) whether distinct patterns of functional organization of the ELN are observable in pwMS compared to matched HCs; (b) whether rsFC in the ELN is associated with language function within the MS group; and (c) whether ELN summary measures are more informative for predicting language function than standard structural and functional MRI measures.

Methods

Participants

Study procedures were approved by Columbia University institutional review board in accordance with ethical guidelines. Written informed consent was obtained from all participants prior to enrollment. For

the MS group, we utilized data collected for MEM CONNECT [20], a prospective cohort of adults diagnosed with relapsing-remitting MS. A separate sample of age, sex, and Intelligent quotient (IQ) matched healthy adults serving in the Reference Ability Neural Networks (RANN) cohort study served as a comparison group [21]. For sample characteristics, see Table 1.

Cognitive measures

All participants completed a comprehensive neuropsychological battery assessing multiple cognitive domains. For this study, we evaluated performance on the following language tests: the Controlled Oral Word Association Test (COWAT): phonemic fluency (FAS) and semantic fluency (Animals); rapid automatized naming: Stroop Word Naming Test, and Stroop Color Naming Test. One-tailed t-tests were used to compare performance across tests based on the expectation that the MS group would show relative decrements compared to the HC group.

MRI data acquisition

In the MS sample, images were acquired on a 3 Tesla MR scanner (GE Discovery) employing the following parameters: Structural images: T1-weighted BRAVO 1 mm sequence, TE/TR=2.7, 7200 ms, voxel resolution=1×1×1mm^3. Functional images: echo planar imaging (EPI), 66 axial slices, TE/TR=25, 850 ms, voxel resolution = 2×2×2 mm^3. During the 9-minute resting-state scan acquisition, participants were instructed to remain still and awake, with eyes closed. In the HC sample, images were acquired on a 3 Tesla MR scanner (Philips Achieva Magnet). Structural images: T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) scan, TE/TR=3, 6500 ms, voxel resolution of 1×1×1 mm^3. Functional images: echo planar imaging, 41 axial slices, TE/TR=20, 2000 ms, voxel resolution = 2×2×2 mm^3. During the 7-minute resting-state scan acquisition, participants were instructed to remain still and awake, with eyes closed [22-24].

Resting-state functional connectivity (rsFC)

Functional connectivity analysis was performed using the CONN toolbox (version 21a; http://www.nitrc.org/projects/conn), implemented in MATLAB 2021a (MathWorks Inc., Natick, MA, USA). Images were preprocessed using CONN toolbox default preprocessing pipeline. Functional data were spatially realigned, unwarped, and slice-time corrected. Outlier scans were identified using conservative thresholds (framewise displacement above 0.5 mm or global blood oxygen-level-dependent (BOLD) signal changes greater than Z = 5). Functional and anatomical images were then normalized into the common stereotaxic Montreal Neurological Institute (MNI) space with 2- and 1-mm isotropic voxels, respectively, and segmented into gray matter, white matter, and cerebrospinal fluid (CSF) tissue classes. Finally, functional data were spatially smoothed using an 8mm full width half maximum (FWHM) Gaussian kernel. Successful normalization and smoothing of functional and anatomical images were confirmed manually for each subject. Next, functional images were denoised using the CONN toolbox default denoising pipeline. Confounding effects were estimated and removed from the BOLD signal for each voxel for each subject using the default anatomical component-based noise correction procedure (aCompCor). Finally, a bandpass filter (0.008, 0.09 Hz) was applied to functional data to investigate low-frequency BOLD signal fluctuations while minimizing influence of physiological and head-movement noise.

rsFC processing

Regions of interest (ROIs) were defined based on the default CONN toolbox atlas, the Schaefer 200-parcel parcellation [25], and the ELN atlas. The default CONN toolbox atlas includes 132 cortical, subcortical, and cerebellar ROIs from the Harvard-Oxford Atlas and Automated anatomical labelling (AAL) atlas 3 [26]. The 17 network Schaefer 200-parcel parcellation was extracted in the 2mm space [27]. A custom atlas was created for the ELN by importing spherical ROIs using the MNI coordinates specified for each region by Tomasi & Volkow [13]. This procedure resulted in a 23 ROI atlas of the ELN (Figure 1). For each participant, Pearson's correlation coefficients were calculated for all possible pairs among the ELN and non-ELN regions in the rest of the brain (Schaefer 17 network 200-parcel parcellation). A 23 × 23 connectivity matrix of Fisher z-transformed r-values for each participant was thus derived for the ELN, and a 23 × 200 connectivity matrix of Fisher z-transformed r-values was derived for the connectivity values between nodes of the ELN and non-ELN nodes. Prior to calculating the ELN summary measures, the matrices were passed through neuroCombat, a site-harmonization tool to reduce scanner effects introduced by differences between the two groups. This method estimates an additive and a multiplicative site-effect coefficient at each parcel, thus accounting for regional scanner differences. neuroCombat has been successfully applied to mitigate scanner differences in previous fMRI studies allowing for harmonization of data collected across multiple scanners [28-30]. The diagonal and negative values of the matrices were set to 0 in the final matrices permitting only positive interactions between ROIs to contribute to derived measures of network interactions.

Calculating ELN summary measures

Four ELN measures were calculated from the resultant matrices: within-ELN connectivity (average connectivity of nodes within the ELN); between-ELN connectivity (average connectivity between nodes of the ELN and nodes of the rest of the brain); Seg-I (relationship of within-ELN connectivity to between-ELN connectivity); and Ant-I (average connectivity of the 11 anterior nodes of the ELN divided by average connectivity of the 12 posterior nodes of the ELN) [19,31]. Higher Seg-I indicates greater 'within-ness' than 'between-ness', that is, greater reliance on within-ELN connections compared to the connections between the ELN and the rest of the brain. Ant-I captures the differences in connectivity of anterior and posterior regions of the ELN. An Ant-I value of 1 indicates equivalent connectivity of anterior and posterior regions, whereas higher Ant-I indicates increasing reliance on anterior regions and lower values indicate greater reliance on posterior regions. See Figure 2 for a graphical representation of the ELN measures.

Structural imaging measures

Additional imaging measures were extracted from structural T1 images of the MS cohort. Cortical thickness of 68 cortical regions (34 per hemisphere) was calculated on lesion in-painted 3D-T1 images using FreeSurfer (V-6.0) with default settings [32].

Diffusion weighted imaging measures

Raw diffusion-weighted images were corrected for distortions caused by motion, eddy current, and field inhomogeneity using FMRIB's Diffusion Toolbox within FSL 6.0.4. Then, probabilistic distribution of 18 major diffusion weighted white matter tracts (corpus callosum-forceps minor, corpus callosum-forceps major, left and right anterior thalamic radiations, uncinate fasciculus, inferior longitudinal fasciculus, cingulum-angular bundle, superior longitudinal fasciculus- temporal segment, superior longitudinal fasciculus-parietal segment, corticospinal tract, and cingulum-cingulate gyrus bundle) in each participant

were extracted using Free Surfer V-6.0, TRACULA [33]. Average fractional anisotropy (FA) and mean diffusivity (MD) for each tract were calculated.

Statistical analyses

For our primary analysis, statistics were conducted with the scipy.stats package in Python.

Group differences in language test performance

To assess whether there were significant differences in language performance, one-tailed t-tests were used to compare the MS to HC group for each of four language tests.

Group differences in ELN summary measures

To assess whether ELN summary measures were differentially expressed between MS and HC groups, two-tailed t-tests were conducted for each of the four summary measures (within-ELN, between-ELN, segregation index, anteriority index).

Relationship between ELN summary measures and performance on language tests

Pearson's correlation coefficients were computed within each diagnostic group to determine whether there were any relationships between summary measures and language performance. In a planned exploratory analysis, we compared performance of ELN summary measures to traditional functional and structural connectivity measures in predicting performance on language tests.

Group differences in pairwise ELN connectivity

To assess whether there were any significant differences in pairwise ELN connections between diagnostic groups, two-tailed t-tests were conducted for each node-node connection with the ELN, as well as each node-node connection from the ELN to the rest of the brain (Schaefer 200-parcel parcellation). All p-values were FDR corrected for multiple comparisons.

Relationship between all language-specific imaging measures and language performance in MS

To compare predictive power of ELN summary measures to other imaging measures, a two-step regression process was applied: First, prior to conducting the regression, the feature set was determined using a data- and empirically driven approach. The feature set for the regression included 4 ELN summary measures, 7 rsFC connections deemed specific to the MS group from the pairwise ELN connectivity analysis, 18 cortical thickness measures from putative language regions in addition to regions that differed at the rsFC level, and 16 diffusion tensor imaging measures (mean diffusivity and fractional anisotropy) of putative language pathways [34]. We calculated mean functional connectivity, cortical thickness, fractional anisotropy, and mean diffusivity to include for reference. Sex and age were also included as features in the regression to rule out their confounding effects. Next, of the 54 pwMS, 9 subjects with missing features were removed from the dataset. This resulted in a 45-subject by 51-feature dataset. Given the suboptimal sample size-to-feature ratio and potential multicollinearity across similar features, 5-fold ridge regression was used to select the top tertile of features. Ridge regression was conducted for each language test individually and the top tertile of features were retained based on the absolute value of the regression feature coefficients [35]. Then, a standard Ordinary Least Squares (OLS) regression was conducted with the top tertile of imaging features to determine the R^2 and p-value of the regression. Finally, by fitting OLS regressors for each language test, we were able to ascertain which

language tests could be significantly predicted by the top tertile of imaging features. If the OLS regression reached significance, the relative coefficient values of the top tertile of features could be analyzed as (weak) representations of feature importance. Ridge regression was performed with sci-kit learn RidgeCV package and OLS regression was performed with statsmodels package in Python.

Results

Group differences in cognitive test performance

The MS group performed significantly worse than HCs on semantic fluency ($p < 0.005$), Stroop Color Naming ($p < 0.005$), and Stroop Word Naming ($p < 0.001$). See Table 2 for full behavioral results.

Group differences in ELN summary measures

The MS group showed higher within-ELN connectivity compared to the HC group ($p < 0.05$). The MS group also demonstrated marginally higher Seg-I compared to HCs, although this did not reach the level of statistical significance ($p = 0.07$). No group differences were found for between-ELN connectivity or Ant-I Figure 3).

Relationship of ELN summary measures to language function

Pearson's correlation analysis revealed no significant relationships between ELN summary measures and language function in either the MS or HC group. However, in the MS group there were several relationships trending towards significance on the Stroop Word Naming test. Both within- and between-ELN connectivity showed trend-level positive correlations with performance on this test ($r = 0.22$, $p = 0.11$; $r = 0.24$, $p = 0.08$, respectively). These trends are reported in reference to a p-value of 0.10, given the small sample size of our study, which may underpower the observed effects [36]. No trend-level relationships were observed between ELN summary measures and any language test in the HC group.

Group differences in pairwise functional connectivity

For exploratory purposes, we evaluated group differences in rsFC among all ELN connections. Five pairwise connections differed for connections *within* the ELN (Figure 4), and two pairwise connections differed for connections *between* nodes of the ELN and nodes of the rest of the brain. While the above connections did not withstand FDR-correction for multiple comparisons, they were included in further exploratory analysis due to their potential relevance as MS-specific language connections.

Relationship of all language-specific imaging measures to language function

Ridge regression models were trained to predict performance on each language test using all ELN, rsFC, diffusion tensor imaging (DTI), and cortical thickness language-specific imaging features. The ridge regression models were trained with a 5-fold cross validation procedure and final R^2 values were calculated for the full training set. With all features included, all four ridge regression models performed with an R^2 value above 0.90. Due to the sample size of the dataset, we were unable to test the ridge regression models on a held-out sample, thus the resultant R^2 values should be cautiously interpreted outside the context of the present sample.

Relationship of key language-specific imaging measures to language function

After retaining the top tertile of features from the ridge regression model of each language test, OLS regression models were fit to predict performance on each language test. Of the four tests, the Stroop

Color Naming test was the only test significantly predicted by the top tertile set of 17 language-specific multimodal imaging features ($R^2 = 0.58$, R^2 corrected = 0.31, $p < 0.05$). The top tertile of features comprised of 4 rsFC measures (including 2 ELN measures), 5 cortical thickness measures, and 8 DTI measures. Age, sex, mean global rsFC, mean cortical thickness, and mean DTI were not retained in the top tertile of features. The relative coefficients (feature weights) of the ridge regression model are shown in Table 3.

Discussion

The main findings of our study point to decrements in language function and alterations within the ELN of pwMS that have been largely overlooked as components of the cognitive profile of pwMS, which has primarily focused on memory and processing speed impairment. In our approach, we derived summary measures to capture large-scale organizational shifts of the ELN. This approach has been used in prior work to test functionally meaningful shifts in memory subnetwork organization [19]. By comparing the ELN summary measures of pwMS to HCs, we found that pwMS exhibited greater within-ELN connectivity relative to HCs. A trend-level difference was also observed for Seg-I. While there were no significant associations of index scores to language function, trend-level associations suggest that we may have been underpowered to detect significant relationships. Finally, our exploratory analysis tested a multimodal model of language function including functional and structural MRI variables, which highlighted the importance of key white matter tracts for predicting language function in pwMS.

Prior studies employing rsFC have generally evaluated connectivity across primary brain networks, e.g., the default-mode network, the salience network, and their constituent nodes [37,38]. Here, we employed a strategy of calculating network-level summary index scores, consistent with our prior work [18]. The advantage of summary measures is that they permit explicit tests of potential mechanisms of large-scale network reorganization to explain dysfunction within a prespecified cognitive domain. Within-ELN connectivity characterizes the intrinsic wiring of the language network, with higher values suggesting stronger 'local' processing. Between-ELN connectivity, conversely, captures the affinity for nodes of the ELN to functionally wire with non-ELN cortical regions, a proxy for 'global' connectivity of the ELN. Seg-I describes the balance between 'local' and 'global' connectivity of the ELN (Figure 2). Comparing these measures between diagnostic groups as well as relating their values to language test performance can thus provide insight into potentially MS-specific neural reorganization related to language dysfunction. Our results highlight higher within-ELN connectivity in the MS group compared to HCs, suggesting stronger connectivity of nodes within the ELN. We also observed slightly elevated (though non-statistically significant) segregation in the MS group pointing toward more local than global connectivity of the ELN.

There is moderate agreement of these results with prior work reporting patterns of network segregation within functional subnetworks in cognitively impaired pwMS [37-39]. Segregated processing is hypothesized to represent functional rerouting as a compensatory mechanism to preserve communication between distant and potentially structurally disconnected brain regions [40]. Some studies, however, have shown alternate patterns of network segregation in pwMS, potentially due to the high sensitivity of the measure to disease stage [41]. Nonetheless, these studies implicate functional rerouting as a probable phenomenon with relevance for cognitive status throughout the MS disease course.

The positive trend of stronger within-ELN connectivity to better language performance supports the hypothesis that functional rerouting is compensatory in early stages of MS disease [42]. As a mechanistic hypothesis, these results point to the possibility that as lesion load and white matter structural damage increase, compensatory functional reorganization takes place [43], leading to greater within-ELN connectivity to maintain language function. In contrast, no relationships were observed between any ELN summary measures and language performance in HCs, suggesting that functional rerouting of large-scale networks may be a marker of compensation in the face of pathological brain changes. Longitudinal studies relating change in functional reorganization to change in cognition across diagnostic groups could help validate this hypothesis.

In our secondary analysis, we compared the predictive value of derived ELN summary measures of rsFC to structural imaging measures. This analysis was conducted in the MS group only, given our aim to explore disease-specific mechanisms of language (dys)function. A subset of the 17 most important imaging features (identified with a data-driven approach; see Table 3) significantly predicted performance on the Stroop Color Naming test. The top feature set was largely dominated by FA and MD of candidate language pathways (superior longitudinal fasciculus and inferior longitudinal fasciculus), consistent with the known importance of demyelination as a hallmark of MS disease and links of white matter microstructure to cognitive impairment [33,41]. The two ELN measures retained in the top feature set were between-ELN connectivity and Ant-I, despite their failure to discriminate between MS and HC groups. The diversity of features retained in the top feature set suggests that brain network rerouting is a complex process that likely occurs structurally, functionally, and at varying levels (i.e., node-node and network-level; Figure 4).

Another advantage of employing network summary measures as opposed to node-node connections in our rsFC analysis is that they minimize individual inhomogeneities (e.g., global rsFC, lateralization differences) that hinder standardization of neuroimaging metrics for use as clinical trial outcomes and in mechanistic models of cognitive impairment [15]. Seg-I, for example, is more resistant to scanner effects as it compares relative differences in network activations on a within-subject basis. If global rsFC was higher overall in one group, Seg-I would be unaffected as the value is dependent on the relative ratio of within-ELN to between-ELN connectivity. Other measures such as within and between-ELN connectivity minimize other inhomogeneities such as lateralization differences across individuals. These measures, calculated downstream of our neuroCombat harmonization, rely on the connectivity of many nodes in a network reducing the potential effects of individual node-node outliers. Finally, network summary measures are replicable, simple to calculate, and theoretically driven making them uniquely useful as mechanistic descriptors of language impairment. In all, our study was well-equipped to address our aim of providing an initial network-level model of language in a sample of adults with MS that would benefit from replication in a larger sample collected across many scanners, to bolster validity of our results.

Language function has largely been omitted from the widely accepted conceptualization of cognitive impairment in MS as dominated by memory and information processing speed dysfunction. Based on growing evidence supporting the prevalence of language dysfunction [6-9], this long-standing oversight needs to be corrected. The present study aims to shift the field's focus to language decrements in MS, and sheds light on a network-level model to guide mechanistic understanding of how language function is disrupted in a 'dysconnection syndrome' [44]. Revisiting the seminal work of Rao and colleagues to

characterize cognitive impairment in MS reveals that in fact verbal fluency was recognized among the top most impaired domains [10]. It is of note that the fluency task they administered was grouped with memory measures, which may have been one factor that set us on a path to disregard the important role of language in MS.

There are some notable limitations to this study. The MS sample primarily comprised patients who were relatively early in disease progression. Future work involving patients in later stages of MS is warranted, as it would extend our findings beyond a relatively limited snapshot of MS disease stages. This could be accompanied by longitudinal studies that relate changes in imaging measures to changes in cognition, which would be a more valid approach for elucidating mechanisms. The sample-size of this study may have limited our ability to detect significant relationships, specifically in the regression analysis. The growing commitment to data sharing and open science in the neuroscience community will hopefully provide the opportunity for replication and follow-ups in a more adequately powered sample of pwMS. A larger cohort would also help address scanner differences between cohorts. In this study, the two cohorts used were collected on different scanners. To harmonize them, we applied neuroCombat, which estimates voxel or parcel-level adjustments to eliminate effects directly associated with different scanner types and protocols. Although the method has been validated for use on a limited number of different scanners, utilizing several cohorts collected across more scanners would yield more effective adjustments based on scanner effects [37].

Conclusions

The results of this proof-of-concept study support the need for future explorations into the neural substrates of language dysfunction in MS. The derived functional network-level measures in addition to the identified structural neuroimaging metrics provide a detailed set of imaging features that can be tested as clinically meaningful predictors and mechanisms of language dysfunction. The proposed framework further facilitates a shift toward the use of standardized network-level neuroimaging metrics in combination with well-defined measures of neuroanatomical language regions to develop a more complete and neuroanatomically specific model of language impairment in MS. With these tools in hand, there is potential for critical advancements into the mechanism and treatment of language impairment in MS.

Acknowledgment: The authors thank Lauren Heuer for assistance in preparing the manuscript for publication, and thank the individuals who participated in the study.

Study Funding: United States Department of Defense Congressionally Directed Medical Research Program (W81XWH-20-1-0503).

Disclosure:

VML has been compensated for advisory or consulting services by the following entities in the last year: Novartis, Biogen. CSR has been compensated for advisory or consulting services by the following entities in the last year: EMD Serono, TG Therapeutics, Horizon, Novartis, Viracta, Genentech. ASR has nothing to disclose. JDD has nothing to disclose. KB has nothing to disclose.LS has nothing to disclose.

References

1. Ciccarelli O, Barkhof F, Bodini B, et al. Pathogenesis of multiple sclerosis: insights from molecular and metabolic imaging. *Lancet Neurol*. 2014;13(8):807-822.
2. Sumowski JF, Benedict R, Enzinger C, et al. Cognition in multiple sclerosis: State of the field and priorities for the future. *Neurology*. 2018;90(6):278-288.
3. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. *Lancet Neurol*. 2008;7(12):1139-1151.
4. Benedict RH, Zivadinov R. Risk factors for and management of cognitive dysfunction in multiple sclerosis. *Nat Rev Neurol*. 2011;7(6):332-342.
5. Benedict RH, Amato MP, DeLuca J, Geurts JJ. Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues. *Lancet Neurol*. 2020;19(10):860-871.
6. Šubert M, Novotný M, Tykalová T, et al. Lexical and syntactic deficits analyzed via automated natural language processing: the new monitoring tool in multiple sclerosis. *Ther Adv Neurol Disord*. 2023;16:17562864231180719.
7. El-Wahsh S, Ballard K, Kumfor F, Bogaardt H. Prevalence of self-reported language impairment in multiple sclerosis and the association with health-related quality of life: An international survey study. *Mul Scler Relat Disord*. 2020;39:101896.
8. Renauld S, Mohamed-Saïd L, Macoir J. Language disorders in multiple sclerosis: A systematic review. *Mult Scler Relat Disord*. Nov 2016;10:103-111. doi:10.1016/j.msard.2016.09.005
9. Lebkuecher AL, Chiaravalloti ND, Strober LB. The role of language ability in verbal fluency of individuals with multiple sclerosis. *Mul Scler Relat Disord*. 2021;50:102846.
10. Rao SM, Leo GJ, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis.: I. Frequency, patterns, and prediction. *Neurology*. 1991;41(5):685-691.
11. Brandstadter R, Fabian M, Leavitt VM, et al. Word-finding difficulty is a prevalent disease-related deficit in early multiple sclerosis. *Mult Scler*. 2020;26(13):1752-1764.
12. Tomasi D, Volkow ND. Network connectivity predicts language processing in healthy adults. *Hum Brain Mapp*. 2020;41(13):3696-3708.
13. Tomasi D, Volkow ND. Resting functional connectivity of language networks: characterization and reproducibility. *Mol Psychiatry*. 2012;17(8):841-854.
14. Malik-Moraleda S, Ayyash D, Gallée J, et al. An investigation across 45 languages and 12 language families reveals a universal language network. *Nat Neurosci*. 2022;25(8):1014-1019.
15. Mohanty R, Gonzalez-Burgos L, Diaz-Flores L, et al. Functional connectivity and compensation of phonemic fluency in aging. *Front Aging Neurosci*. 2021;13:644611.
16. He X, Bassett DS, Chaitanya G, Sperling MR, Kozlowski L, Tracy JI. Disrupted dynamic network reconfiguration of the language system in temporal lobe epilepsy. *Brain*. 2018;141(5):1375-1389.
17. Blecher T, Miron S, Schneider GG, Achiron A, Ben-Shachar M. Association between white matter microstructure and verbal fluency in patients with multiple sclerosis. *Front Psychol*. 2019;10:1607.
18. Leavitt VM, Dworkin JD, Buyukturkoglu K, Riley CS, Ritchey M. Summary metrics of memory subnetwork functional connectivity alterations in multiple sclerosis. *Mul Scler*. 2022;28(12):1963-1972.
19. Stern Y, Habeck C, Steffener J, et al. The Reference Ability Neural Network Study: motivation, design, and initial feasibility analyses. *Neuroimage*. 2014;103:139-151.
20. Buyukturkoglu K, Vergara C, Fuentealba V, et al. Machine learning to investigate superficial white matter integrity in early multiple sclerosis. *J Neuroimaging*. Jan 2022;32(1):36-47. doi:10.1111/jon.12934
21. Argiris G, Stern Y, Habeck C. Age-related disintegration in functional connectivity: Evidence from Reference Ability Neural Network (RANN) cohort. *Neuropsychologia*. Jun 18 2021;156:107856. doi:10.1016/j.neuropsychologia.2021.107856

22. Stern Y, Gazes Y, Razlighi Q, Steffener J, Habeck C. A task-invariant cognitive reserve network. *Neuroimage*. 2018;178:36-45.
23. Ryu H, Habeck C, Stern Y, Lee S. Persistent homology-based functional connectivity and its association with cognitive ability: Life-span study. *Hum Brain Mapp*. 2023;
24. Whitfield-Gabrieli S, Nieto-Castanon A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect*. 2012;2(3):125-41. doi:10.1089/brain.2012.0073
25. Lawrence RM, Bridgeford EW, Myers PE, et al. Standardizing human brain parcellations. *Scientific Data*. 2021/03/08 2021;8(1):78. doi:10.1038/s41597-021-00849-3
26. Schaefer A, Kong R, Gordon EM, et al. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex*. Sep 1 2018;28(9):3095-3114. doi:10.1093/cercor/bhx179
27. Fortin JP, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. Feb 15 2018;167:104-120. doi:10.1016/j.neuroimage.2017.11.024
28. Cetin Karayumak S, Bouix S, Ning L, et al. Retrospective harmonization of multi-site diffusion MRI data acquired with different acquisition parameters. *Neuroimage*. Jan 1 2019;184:180-200. doi:10.1016/j.neuroimage.2018.08.073
29. Yu M, Linn KA, Cook PA, et al. Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Hum Brain Mapp*. Nov 2018;39(11):4213-4227. doi:10.1002/hbm.24241
30. Chan MY, Park DC, Savalia NK, Petersen SE, Wig GS. Decreased segregation of brain systems across the healthy adult lifespan. *Proc Natl Acad Sci U S A*. Nov 18 2014;111(46):E4997-5006. doi:10.1073/pnas.1415122111
31. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*. Feb 1999;9(2):179-94. doi:10.1006/nimg.1998.0395
32. Yendiki A, Panneck P, Srinivasan P, et al. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Front Neuroinform*. 2011;5:23. doi:10.3389/fninf.2011.00023
33. Houston J, Allendorfer J, Nenert R, Goodman AM, Szaflarski JP. White Matter Language Pathways and Language Performance in Healthy Adults Across Ages. *Front Neurosci*. 2019;13:1185. doi:10.3389/fnins.2019.01185
34. Schoonheim MM, Geurts J, Wiebenga OT, et al. Changes in functional network centrality underlie cognitive dysfunction and physical disability in multiple sclerosis. *Mult Scler*. Jul 2014;20(8):1058-65. doi:10.1177/1352458513516892
35. Cule E, Vineis P, De Iorio M. Significance testing in ridge regression for genetic data. *BMC Bioinformatics*. Sep 19 2011;12:372. doi:10.1186/1471-2105-12-372
36. Jewell T. Library Guides: AMA Style Guide: In-Text Citations. 2011;
37. Schoonheim MM, Broeders TAA, Geurts JJG. The network collapse in multiple sclerosis: An overview of novel concepts to address disease dynamics. *Neuroimage Clin*. 2022;35:103108. doi:10.1016/j.nicl.2022.103108
38. Eijlers AJ, Meijer KA, Wassenaar TM, et al. Increased default-mode network centrality in cognitively impaired multiple sclerosis patients. *Neurology*. Mar 7 2017;88(10):952-960. doi:10.1212/wnl.0000000000003689
39. Welton T, Constantinescu CS, Auer DP, Dineen RA. Graph Theoretic Analysis of Brain Connectomics in Multiple Sclerosis: Reliability and Relationship with Cognition. *Brain Connect*. Mar 2020;10(2):95-104. doi:10.1089/brain.2019.0717
40. Liu Y, Wang H, Duan Y, et al. Functional Brain Network Alterations in Clinically Isolated Syndrome and Multiple Sclerosis: A Graph-based Connectome Study. *Radiology*. Feb 2017;282(2):534-541. doi:10.1148/radiol.2016152843
41. Roberts RE, Anderson EJ, Husain M. White matter microstructure and cognitive function. *Neuroscientist*. Feb 2013;19(1):8-15. doi:10.1177/1073858411421218

42. Hawellek DJ, Hipp JF, Lewis CM, Corbetta M, Engel AK. Increased functional connectivity indicates the severity of cognitive impairment in multiple sclerosis. *Proc Natl Acad Sci*. 2011;108(47):19066-19071.
43. Castellazzi G, Debernard L, Melzer TR, et al. Functional connectivity alterations reveal complex mechanisms based on clinical and radiological status in mild relapsing remitting multiple sclerosis. *Front Neurol*. 2018;9:690.
44. Sjøgård M, Wens V, Van Schependom J, et al. Brain dysconnectivity relates to disability and cognitive impairment in multiple sclerosis. *Hum Brain Mapp*. Feb 15 2021;42(3):626-643. doi:10.1002/hbm.25247

Table 1. General Characteristics of the study group.

Variables	MS (N=54)	HCs (N=54)	p-value
Sex (N%)			
Men	12 (22.2)	14 (25.9)	0.65
Women	42 (77.8)	40 (74.1)	
Age	39.8 ± 9.8^	39.0 ± 9.8	0.73
Estimated intelligence quotient [a]	111 ± 7.6	111 ± 8.6	0.64

[a]As estimated by Wechsler Test of Adult Reading, ^ mean ± SD.
MS: multiple sclerosis; HCs: Healthy controls.

Table 2. Language test performance for study groups.

Variables	MS	HCs	t-score	p-value
Verbal fluency				
FAS (phonemic fluency)	42.2±13.7^	44.3±10.1	-0.88	0.19
Animals (semantic fluency)	22.0±6.1	24.6±4.9	-2.45	0.008*
Rapid automatized naming				
Stroop Color	71.5±11.9	78.1±14.1	-2.65	0.005*
Stroop Word	95.1±16.7	105.5±16.8	-3.20	0.0009**

^mean ± SD; *p<0.01, **p<0.001.

MS: multiple sclerosis; HCs: Healthy controls.

Table 3. Ridge regression coefficients for Stroop color test.

Feature	Weight
LH SLFP MD	+11.349
LH SLFT FA	+6.547
RH SLFT FA	+6.311
Left inferior temporal-right inferior temporal	+5.416
LH ILF FA	+5.292
ELN between-ness	+4.002
RH inferiorparietal thickness	+3.753
Broca's area-left IPFC	+3.738
LH SLFT MD	+3.564
anteriority index	-3.773
LH superiortemporal thickness	-3.862
RH parsorbitalis thickness	-4.435
RH SLFP FA	-5.660
LH ILF MD	-5.896
LH precuneus thickness	-7.363
RH SLFT MD	-7.504

FA: fractional anisotropy; MD: mean diffusivity; LH: left; RH: right; SLFP: superior longitudinal fasciculus-parietal endings; SLFT: superior longitudinal fasciculus-temporal endings; ILF: inferior longitudinal fasciculus.

Legends

Figure 1. Nodes and connections of the ELN defined by Tomasi & Volkow, 2012[13].
(a) axial view, (b) coronal view, (c) left sagittal view. 1= Wernicke's area; 2= Right inferior parietal; 3= Middle frontal; 4= Pars opercularis; 5= Left pars orbitalis; 6= Right pars orbitalis; 7= Left inferior temporal; 8= Right inferior temporal; 9= Superior frontal; 10= Cerebellum; 11= Broca's area; 12= Pars triangularis; 13= Left caudate; 14= Right caudate; 15= Putamen/globus pallidus; 16= Ventral thalamus; 17= Striate; 18= Extrastriate; 19= Posterior parietal; 20= Superior Parietal; 21= Left superior temporal; 22= Right superior temporal; and 23= Cingulate.

Figure 2. Derivation of ELN summary measures.
(a) Square matrix of rsFC within the 23 ELN nodes used to calculate within-ELN connectivity. Left sagittal view of within-ELN connectivity. (b) Matrix of rsFC between 23 ELN ROIs and 200 non-ELN ROIs of the rest of the brain used to calculate between-ELN connectivity. Darker red edges represent stronger connectivity between nodes (higher r-values) while lighter red and light blue represent low connectivity (see scale in *a*). Negative connectivity values are set to zero in the calculation of ELN summary measures. (c) Anterior and posterior nodes of the ELN. Dotted line is a reference for anterior versus posterior regions of the ELN. (d) Segregation index equation. ELN: extended language network.

Figure 3. Group differences in network summary metrics.
Quantification of (a) within-ELN connectivity, (b) between-ELN connectivity, (c) Seg-I, and (d) Ant-I for individuals with MS and HCs. Means are marked with "X". *$p < 0.05$. ELN: extended language network; Se-I: segregation index; Ant-I: anteriority index; MS: Multiples sclerosis; HCs: Healthy controls.

Figure 4. Group average within-ELN connectivity.
(a, b) Average within-ELN connectivity in patients with MS and HCs. Red lines represent the fisher z-transformed r values of connectivity between nodes with darker lines representing stronger connectivity (i.e., higher r-values). (c) Group differences in within-ELN connectivity thresholded at $p < 0.05$ uncorrected. Connections include middle frontal to right pars orbitalis, left pars orbitalis to pars triangularis, right pars orbitalis to right inferior temporal, left inferior temporal to right inferior temporal, and Broca's area to pars triangularis. ELN: extended language network; MS: Multiples sclerosis; HCs: Healthy controls.

Figure 1

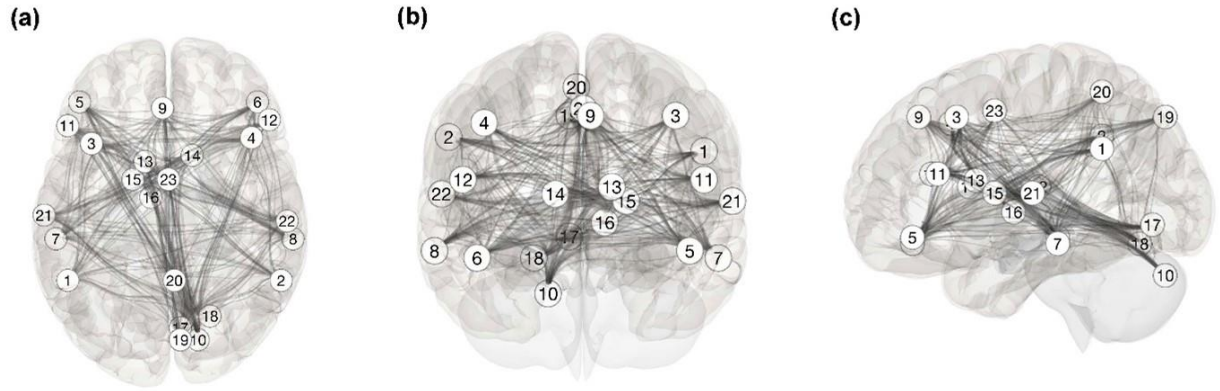


Figure 2

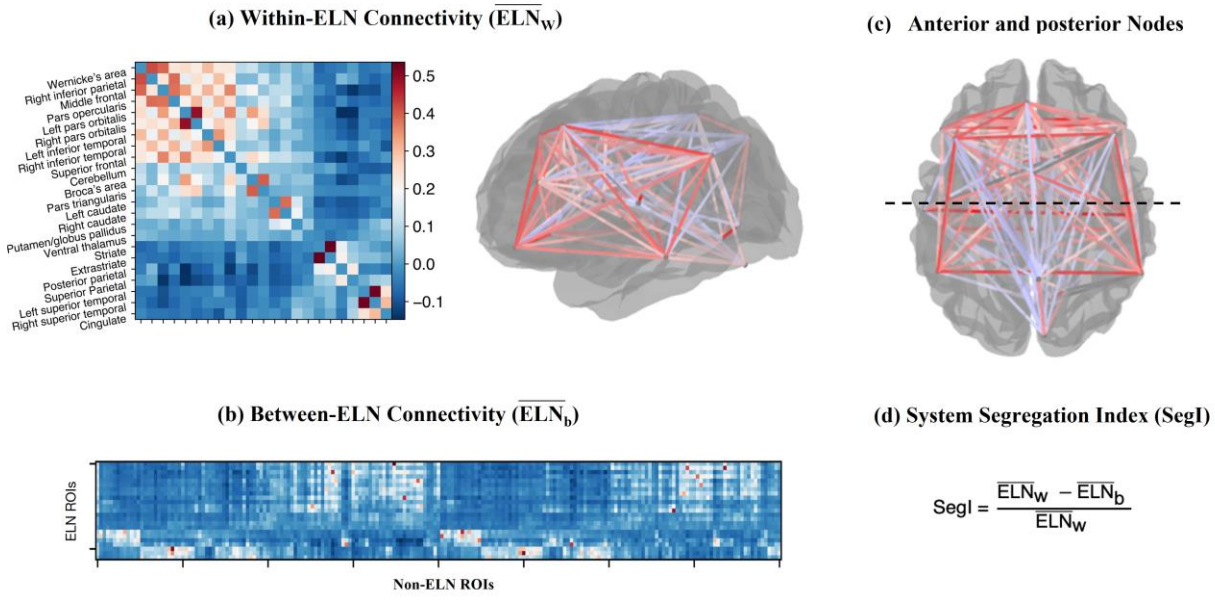


Figure 3

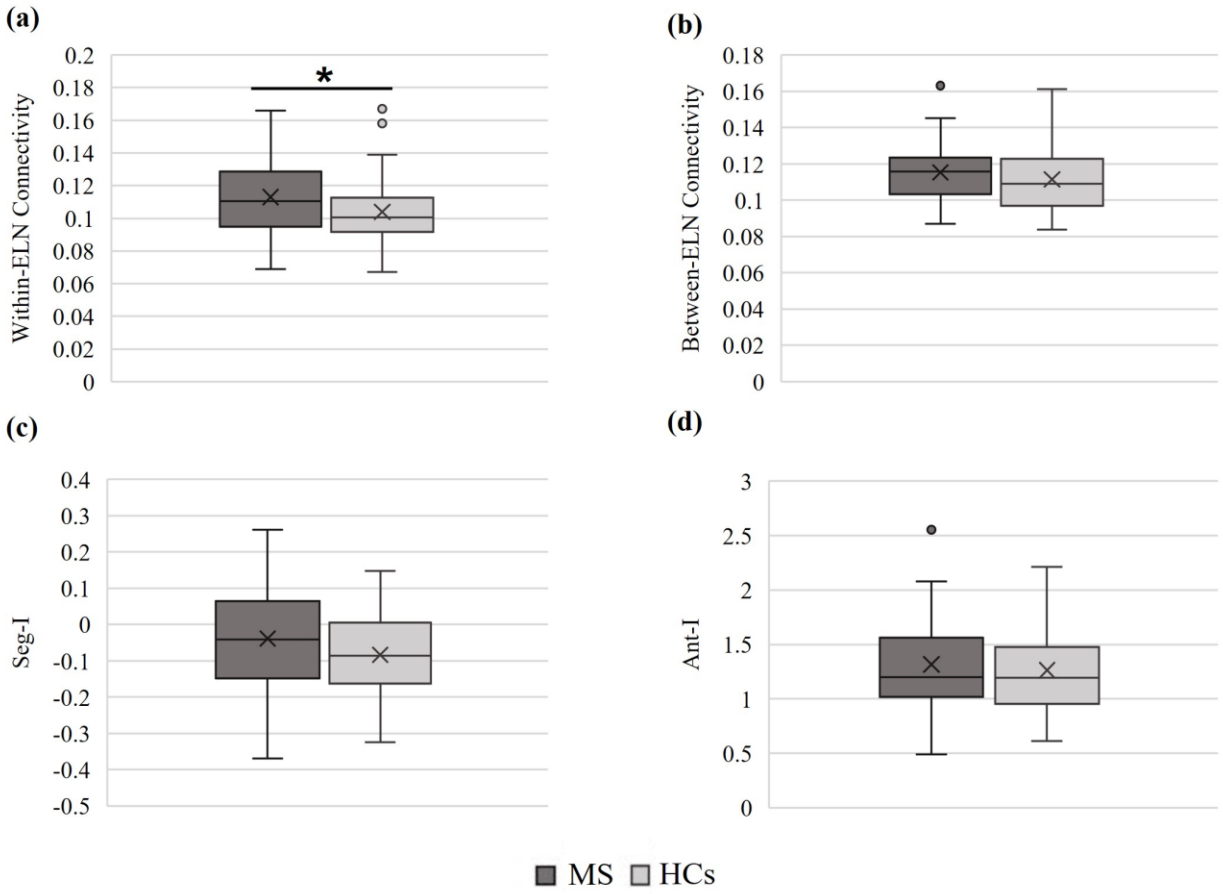
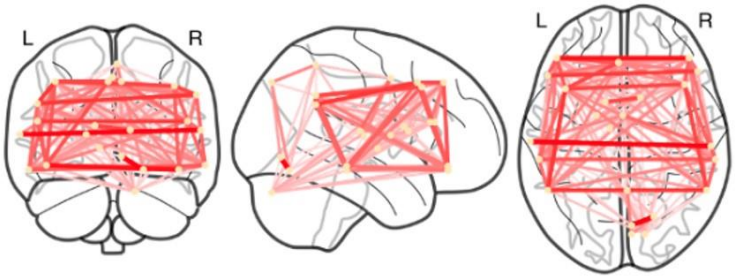
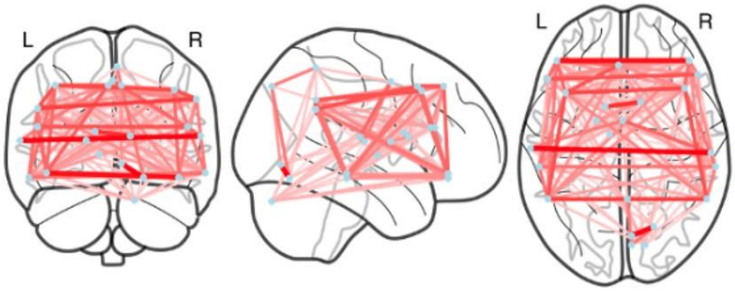


Figure 4

(a) Average within-ELN connectivity in MS



(b) Average within-ELN connectivity in HCs



(c) Difference in average within-ELN connectivity in (MS > HCs)