MONASH University

Speech Perception in Noise in Multiple Sclerosis

Pippa Lorena Iva
Bachelor of Behavioural Neuroscience (Hons)

Supervisors: Professor Ramesh Rajan, Associate Professor Joanne Fielding, Associate Professor Anneke van der Walt, Dr Russell Martin

A thesis submitted for the degree of *Doctor of Philosophy* at
Monash University in 2020

Monash Biomedicine Discovery Institute
Department of Physiology

Table of Contents

i Pretext **7**
ii Thesis including published works declaration **8**
iii Acknowledgements **10**
iv Main Abbreviations **11**
1 Chapter One: Auditory Processing in Multiple Sclerosis **12**
1.1 Introduction to Literature Review **12**
1.2 Nature of MS **13**
1.2.1 Pathophysiology 13
1.2.2 Diagnosis 13
1.2.3 Clinical subtypes 14
1.2.4 Motor symptoms 15
1.2.5 Cognition 16
1.2.6 Overview 17
1.3 Hearing sensitivity and auditory flow in MS **17**
1.3.1 Pure tone thresholds 17
1.3.2 Acute influences on pure tone thresholds 18
1.3.3 Chronic influences on pure tone thresholds 19
1.4 Central auditory processing deficits **21**
1.5 Auditory flow in MS: Auditory evoked potentials **22**
1.5.1 Brainstem auditory evoked potentials 22
1.5.2 Auditory middle latency response 23
1.5.3 Cognitive potential (P300) 23
1.6 Processing cues to auditory space **24**
1.6.1 Lateralisation tests 24
1.6.2 Relationship between lateralisation performance and MRI lesions 25
1.6.3 Interaural discrimination tests 25
1.6.4 Relationship between interaural discrimination and brainstem auditory responses 26
1.7 Temporal processing in MS **27**
1.7.1 Masking level differences 27
1.7.2 Frequency pattern test 28
1.7.3 Gaps in noise test 29
1.8 Speech perception in MS **30**
1.8.1 Speech in silence 30
1.8.2 Dichotic speech paradigms 31
1.8.3 Speech in noise paradigms 32
1.9 Conclusion **33**
2 Chapter Two: Speech Discrimination Impairments as a Marker of Disease Severity in Multiple Sclerosis **36**
2.1 Introduction **38**

2.2 Methods 39
2.2.1 Participants 39
2.2.2 Study Overview 39
2.2.3 Pure-tone audiometry 39
2.2.4 Speech-in-noise discrimination tasks 39
2.2.5 Auditory attention and discomfort questionnaire 40
2.2.6 Generalised linear mixed model 40
2.3 Results 41
2.3.1 Participant groups 41
2.3.2 Speech discrimination in noise 42
2.3.3 Identification of sentences in speech-weighted noise 42
2.3.4 Identification of sentences in multi-talker babble 42
2.3.5 Estimating psychometric functions 44
2.3.6 Modelling the factors that impact on SiN discrimination 44
2.3.7 Disease factors that impact SiN discrimination 47
2.3.8 Noise and daily life events 49
2.4 Discussion 50
3 Chapter Three: Speech Discrimination Tasks: A Sensitive Sensory and Cognitive Measure in Early and Mild Multiple Sclerosis 54
3.1 Introduction 55
3.2 Methods 58
3.2.1 Participants 58
3.2.2 Neuropsychological testing 58
3.2.3 Pure-tone audiometry 58
3.2.4 Auditory attention and discomfort questionnaire (AADQ) 59
3.2.5 Loudness sensitivity test 59
3.2.6 Speech in noise discrimination tasks 59
3.2.7 Statistical analyses 60
3.3 Results 62
3.3.1 Participant demographics, characteristics, and audiometric hearing status 62
3.3.2 Sensory discomfort to high intensitity stimuli 67
3.3.3 Discrimination of sentences in noise 69
3.3.4 Discrimination of keywords in noise 71
3.3.5 Words in babble discrimination 74
3.3.6 Correlations of speech discrimination measures to standard neuropsychological tests 77
3.3.7 Comparison of all SiN measures to discriminate between controls and all MS group 80
3.3.8 Classification value of speech discrimination tasks to classify all mild (early and late) pwMS from controls 82
3.4 Discussion 87
4 Chapter Four: Discriminating Spatialised Speech in Complex Environments in Multiple Sclerosis 93

4.1 Introduction 95
4.2 Methods 97
4.2.1 Participants 97
4.2.2 Neuropsychological testing 97
4.2.3 Audiometry 97
4.2.4 Interaural time differences 97
4.2.5 Spatial speech-in-noise task 98
4.2.6 Coordinate response measure task 99
4.2.7 Statistical analyses 101
4.3 Results 102
4.3.1 Participant demographics, characteristics, and audiometric hearing status 102
4.3.2 Discrimination of interaural time differences 103
4.3.3 Spatial release of speech from multi-talker babble noise 104
4.3.4 Spatial release of target speech from masker speech in the coordinate response measure task 106
4.3.5 Multivariate analysis on factors influencing spatial release from masking in the coordinate response measure task 108
4.3.6 Multivariate analysis on disease factors influencing performance at each spatial condition in the coordinate response measure task 110
4.3.7 Errors in the coordinate response measure task 112
4.3.8 Impact of disease severity on discriminating spatialised speech in babble 114
4.3.9 Multivariate analysis on factors influencing performance in the spatialised sentences in babble task ..... .116
4.4 Discussion 119
5 Chapter Five: An Exploratory Study on the Relationship Between Brain Atrophy and Processing Degraded Speech Abilities in Multiple Sclerosis (Preliminary Findings) 124
5.1 Introduction 125
5.2 Methods 128
5.2.1 Participants 128
5.2.2 Neuropsychological testing 128
5.2.3 Magnetic Resonance Imaging 128
5.2.4 Audiometry 128
5.2.5 SiN Discrimination Tasks 128
5.2.6 Statistical analysis 128
5.3 Results 130
5.3.1 Participant demographics, characteristics, and audiometric hearing status 130
5.3.2 Discrimination of sentences in noise 131
5.3.3 Magnetic Resonance Imaging analysis 132
5.4 Discussion 136
6 Chapter Six: General Discussion 139
6.1 Study design aspects 139
6.1.1 EDSS classification groups 140

6.1.2 Sub-group analysis vs. single case reports 140
6.2 Auditory processing in MS 141
6.2.1 Peripheral auditory processing in MS 141
6.2.2 Central auditory processing in MS: Brainstem processing 142
6.2.3 Central auditory processing in MS: Speech-in-noise processing 143
6.2.4 Deficits in psychoacoustic tasks 145
6.2.5 Subjective auditory processing complaints 146
6.3 2. The utility of speech-in-noise tests to evaluate sensory and cognitive deficits in MS 147
6.3.1 Speech-in-noise paradigms in other populations 147
6.3.2 Multi-talker babble: a potent masker for early and mildly affected pwMS 148
6.3.3 Higher order processes required to reconcile perceptual ambiguity 151
6.3.4 Neurodegenerative pathology and its association with speech-in-noise performance 152
6.3.5 Advantages of SiN tasks 153
6.3.6 Limitations of SiN tasks 153
6.4 Concluding remarks 154
7 Supplementary Material 155
8 References 163

Copyright notice

i Pretext

Multiple sclerosis (MS) is a multi-component disease where inflammatory and neurodegenerative processes disrupt wide-ranging cerebral systems, including auditory networks. People with MS (pwMS) rarely present with cochlear hearing loss, but often have abnormal centralised processes which can implicate the processing of complex sounds such as speech; the main form of communication that connects social human beings. Speech processing often takes place in the presence of noise, and requires not just sensory encoding, but higher-order processes such as working memory, attention, emotion, and executive function to reconcile perceptual ambiguity. Despite the impact on a person's ability to navigate the world, build relationships and maintain employability, studies of speech in noise (SiN) perception in pwMS have been limited.

Speech in noise (SiN) perception in pwMS was uniquely evaluated under 'simulated real-world' conditions using a battery of psychoacoustic tests (**Chapter Two**). After a routine audiometric evaluation, controls and pwMS listened to pre-recorded speech in 'ecologically relevant' multi-talker babble and steady speech-weighted noise at varying signal-to-noise ratios (SNRs) and were required to verbally repeat the target speech. Despite normal hearing, MS psychometric discrimination curves which model the relationship between SNR and speech discrimination accuracy did not change in slope (sentences/dB) but shifted to higher SNRs (dB) compared to controls. This suggested that pwMS required louder target signals to achieve the same level of discrimination accuracy as controls. The magnitude of the shift in the curve systematically increased with greater disability, as evaluated by Expanded Disability Status Scale (EDSS) scores. In SiN conditions where target speech emanated from a different spatial location to the noise, pwMS displayed pronounced deficits in SiN perception compared to co-localised conditions (**Chapter Four**). Although pwMS across a spectrum of disability (mild to severe) had impairments in at least one SiN task, only participants with severe disability self-reported significant audio-attentional difficulty during daily life events. The absence of self-reported auditory difficulty is likely to reflect the intrinsic and extrinsic redundancy in auditory processing.

Susceptibility to speech degradation in babble correlated negatively with performance on neuropsychological testing in pwMS, suggesting that poor SiN perception may manifest from cognitive impairment in pwMS (**Chapters Three & Four**). To investigate the pathological underpinnings of SiN perception in pwMS, the relationship between volumetric measures of several key neuroanatomical regions from magnetic resonance imaging (MRI) and SiN perception measures was determined (**Chapter Five**). Preliminary findings suggested poorer SiN perception in babble was associated with smaller temporal white matter volume, but not grey matter.

Neuropsychological measures currently used in the clinic require trained personnel and considerable time to administer, therefore, not all pwMS receive formal cognitive assessment despite cognitive impairment affecting up to 70% of pwMS. To address this clinical need, we suggest SiN tasks be employed as a screening tool for patients who require further cognitive assessment due to their ease of use, speed, cost effectiveness – features which are advantageous in a clinical setting.

ii Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes two submitted publications. The core theme of the thesis is speech in noise perception in multiple sclerosis. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the Department of Physiology, Monash University, under the primary supervision of Prof. Ramesh Rajan, co-supervised by Assoc. Prof. Joanne Fielding, Assoc. Prof. Anneke van der Walt, and Dr. Russell Martin.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research. In the case of Chapters Two and Three, my contribution to the work involved the following:

Thesis Chapter	**Publication Title**	**Status** *(published, in press, accepted or returned for revision, submitted)*	**Nature and % of student contribution**	**Co-author name(s) Nature and % of Co-author's contribution *denotes co-author was a Monash student**
2	Speech Discrimination Impairments as a Marker of Disease Severity in Multiple Sclerosis	Submitted to Multiple Sclerosis and Related Disorders, under review	60%. Input into experimental design, collected data, conducted analysis, and wrote the manuscript	1) Joanne Fielding, input into manuscript 5% 2) Meaghan Clough, input into manuscript 5% 3) Owen White, input into manuscript 5% 4) Gustavo Noffs, contributed to data collection 2.5% 5) Branislava Godic, contributed to data collection 2.5% 6) Russell Martin, input into manuscript 5% 7) Anneke van der Walt, input into manuscript 5% 8) Ramesh Rajan, designed experiment, input into manuscript 10%

3	Speech Discrimination Tasks: A Sensitive Sensory and Cognitive Measure in Early and Mild Multiple Sclerosis	Submitted to Frontiers Neuroscience, under review	60%. Input into experimental design, collected data, conducted analysis, and wrote the manuscript	1) Joanne Fielding, contributed to data collection, input into manuscript 7.5% 2) Meaghan Clough, input into manuscript 5% 3) Owen White, contributed to data collection, input into manuscript 7.5% 4) Branislava Godic, contributed to data collection, input into manuscript 5% 5) Russell Martin, input into manuscript 5% 6) Ramesh Rajan, designed experiment, input into manuscript 10%

These submitted papers have been presented in this order to generate a consistent presentation within the thesis.

Student name: Pippa Iva

Student signature: [redacted] **Date: 11/09/2020**

I hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor name: Ramesh Rajan

Main Supervisor signature: [redacted] **Date: 11/9/2020**

iii Acknowledgements

First and foremost, I would like to extend my sincerest thanks to my supervisor, Prof. Ramesh Rajan, who provided me with immense support and opportunities throughout my research degree. It has been such a privilege to be supervised by someone I have great respect for – your knowledge and guidance has formed the foundation of my career in this field and I will forever be thankful. Your words of encouragement, kindness, and empathy for both professional and personal aspects during my project have really seen me through. It was a pleasure to be taught by you.

I would like to acknowledge the valuable input of my co-supervisors, Assoc. Prof. Joanne Fielding, Assoc. Prof. Anneke van der Walt and Dr Russell Martin. To Assoc. Prof. Joanne Fielding, thank you for always being so welcoming and inclusive of me into your research group and for your expertise, and direction. To Assoc. Prof. Anneke van der Walt, thank you for the wonderful opportunity to work with you and for your expert insight you readily provided whenever I needed. To Dr Russell Martin, I really appreciated your time and patience with all my questions- you always took the time after lab meetings to have in-depth conversations with me about my research.

I would also like to extend my sincerest gratitude to Professor Owen White, Dr Gustavo Noffs and Dr Meaghan Clough for your collaboration and expertise. A special thanks goes to Dr Meaghan Clough – your genuine kindness, generosity with your time, and exceptional knowledge about the field has been so valuable to me throughout the years.

More broadly, many thanks to those in the Department of Physiology for the friendships and support – particularly to lab members, Branislava Godic and Angela Ediriweera. Also, to my wonderful and understanding friends, thank you for the unwavering support and encouragement I needed to achieve this.

My heartfelt thanks to my family, Vas and Emy for your love and support, my sister – and rock – Kylie, who supported me with not just words of encouragement but who would often drop food at my house and offer to take on a lot of my responsibilities to give me more time to write- I'm incredibly grateful to you. To my fiancé, Daniel, who endured this journey with me, supported me every step of the way and always encouraged me to keep going- this would not have been possible without you.

Finally, thank you to the wonderful people who gave up some of their generous time to participate in my experiments. Some of you attended several sessions – without you this thesis would not have been possible.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

iv Main Abbreviations

AADQ: Auditory attention and discomfort questionnaire
BAEP : Brainstem auditory evoked potential
BDI : Beck's depression inventory
BKB : Bamford-Kowl-Bench
BN : Babble noise
CIS : Clinically isolated syndrome
CNS : Central nervous system
CRM : Coordinate response measure
CVLT : California verbal learning test
DST : Digit span test
EDSS : Expanded disability status scale
GIN : Gaps-in-noise
ILD : Interaural level difference
ITD : Interaural time difference
JND : Just noticeable difference
LDL : Loudness discomfort level
MRI : Magnetic resonance imaging
MS : Multiple sclerosis
NART : National adult reading test
PASAT : Paced auditory serial addition test
pwMS : People with multiple sclerosis
RR : Relapsing remitting
SiN : Speech-in-noise
SNR : signal-to-noise ratio
SP : Secondary progressive
SRM : Spatial release from masking
SRT : Speech reception threshold
SSiN : Spatialised speech-in-noise
SWN : Speech-weighted-noise
WiN : Words-in-noise

Chapter One: Auditory Processing in Multiple Sclerosis

1.1 Introduction to Literature Review

Multiple sclerosis (MS) is a chronic, autoimmune disorder of the central nervous system (CNS) and the leading cause of non-traumatic neurological disability to afflict young and middle-aged adults in high income countries (1, 2). Driven by a complex interplay of inflammatory and neurodegenerative changes, the clinical course of MS is varied and unpredictable (3-5). Most people with MS (pwMS) experience discrete 'episodes' of transient or permanent neurological dysfunction as a consequence of immune-mediated inflammation resulting in demyelination and nerve conduction disruption (6). Symptoms during these episodes largely depend on the site of CNS involvement, but often involve sensory, motor, bladder/bowel, cerebellar, brainstem and cognitive changes (7, 8).

MS can impact any CNS neural system, including the auditory system. Successful auditory processing allows sound detection by the external ear up to cortex for the perceptual processing of meaning from the signal; enabling us to socialise, work and communicate. Despite having a direct impact on quality of life, investigations of auditory processing disorders in the context of MS are comparatively less characterized than other sensory symptoms such as hypoesthesia, paresthesia, and visual disturbances. Since the loci of lesions within the CNS can occur at multiple sites along this pathway and develop or regress at different rates between patients with MS, nearly any auditory symptomatology pattern may result (9). The involvement of MS in auditory processing can be detected by magnetic resonance imaging (MRI), auditory evoked potentials (AEP) and psychoacoustic measures (10). Numerous efforts have been made to describe abnormal prevalence of these individual measures in the MS population, as well as elucidate the relationships between them, however not many clear associations have emerged. For example, one MS patient may have normal peripheral hearing sensitivity, speech understanding and discrimination of interaural intensity differences but an abnormal auditory brainstem response (ABR), whereas another may have normal ABR, normal hearing sensitivity, but reduced speech intelligibility. From this, it is established that a complex relationship exists amongst lesion site and auditory performance (10); and such findings provide support for the concept of a multiple test battery to describe auditory deficits in this unique population. Furthermore, the literature in this field is often inconsistent. MS is a heterogenous disease that is variable not only between patients, but also within patients over time. There have been no large-scale studies that evaluate central auditory processing in pwMS; which may explain why study findings can be conflicting at times; the only certainty is that there is an issue.

Psychoacoustic, electrophysiological, and imaging studies are reviewed in an effort to describe MS impairments in audition. This review will highlight the widely-accepted view that the standard pure-tone audiogram and speech reception threshold (SRT) tests alone do not provide a comprehensive assessment of auditory function in MS (11). This clinical population display subtle yet detectable auditory deficits using sensitive, specialized measures involving central rather than peripheral auditory processing (12).

1.2 Nature of MS

Before embarking on the patterns of auditory impairment in MS, a brief review of some of the key aspects of the disease is pertinent. MS is an inflammatory disease of the CNS that affects almost 2,500,000 individuals worldwide (13, 14). MS was first described in 1868 by Jean-Martin Charcot as multifocal destruction of myelinated CNS tracts resulting in a variety of neurological dysfunction (15, 16). The disease course is highly variable but in most cases presents as a pathophysiological process that involves the complex interplay of damage and repair mechanisms which determine the clinical course of the disease (4, 15). There is a large body of evidence to support the idea that MS can be conceptualized as a two-stage disease involving transient episodes of inflammation, often followed by a neurodegenerative component associated with irreversible loss of axons and neurons, leading to progressive disability (3, 4, 17).

1.2.1 Pathophysiology

MS is thought to be of autoimmune origin as myelin-specific T lymphocytes (T cells) activated in the periphery translocate into the CNS to induce inflammation and demyelination (18, 19). Once entering the CNS, the T cells are reactivated by antigen-presenting cells, which present major histocompatibility complex (MHC) class II-associated peptides, resulting in naïve T cells differentiating into various subsets of effector T cells with distinct functions directed at the degradation of myelin and inflammation (20, 21). Simultaneously, B cells are also activated and participate in the development of CNS lesions by secreting injuring autoantibodies, presenting autoantigen to T cells and contributing to the secretion of proinflammatory cytokines (22-24). This cascade of inflammatory responses is accompanied by disturbances of the blood brain barrier to facilitate the attack on myelin and oligodendrocytes (20). The trigger factors of MS are still largely unknown and at present, there are no well-established factors to assist disease prevention (25). The majority view at present is that MS is a complex disorder acquired by genetically susceptible individuals who are exposed to unidentified environmental infectious agents, which subsequently triggers an auto immune attack on the myelin sheath of axons in the CNS (2, 26-28).

1.2.2 Diagnosis

The McDonald diagnostic criteria for MS has been extensively assessed and used for the diagnosis for MS since it was introduced in 2001 (29), revised in 2005 (30) and again in 2010 (31). This criterion allows an accurate diagnosis of MS by utilizing modern diagnostic investigations including MRI, evoked potentials and cerebral spinal fluid (CSF) immunologic changes. The majority of MS diagnostic criteria are based on demonstrating the following:

1) Objective evidence of two CNS lesions separated both in time and space of occurrence and;
2) Other potential causes for the CNS lesions to be ruled out (29-31).

Dissemination in space requires demonstration that the inflammatory-demyelinating injury involves at least two discrete neuro-anatomic areas within the CNS. This can be established using MRI techniques, to identify at least one T2 lesion in at least two out of four areas of the CNS: periventricular, juxtacortical, infratentorial, or spinal cord (29, 31). To provide evidence of dissemination in time, there must be a new lesion on follow up MRI, with reference to a baseline scan, or simultaneous presence of symptomatic and asymptomatic lesions at any time. The purpose of this requirement is to rule out monophasic illnesses being mistaken for MS, which is primarily a reoccurring inflammatory process (29-31). A positive CSF can be determined with the presence of elevated Oligoclonal immunoglobulin G (IgG) bands in CSF or elevated IgG index to affirm the inflammatory demyelinating nature of the underlying condition, and thereby can be helpful in excluding other causes. Although the supportive diagnostic tool of CSF analysis is not formally required for relapsing forms of MS due to the emphasis on MRI findings, CSF cytokine/chemokine profiling is necessary for the differential diagnosis of the primary progressive MS phenotype, the least common clinical phenotype of the disease (32, 33).

1.2.3 Clinical subtypes

The first episode suggestive of a demyelinating attack is referred to as a clinically isolated syndrome (CIS) and is often a precursor of MS (8, 34, 35). Not all individuals who experience CIS will go on to develop MS, but the risk is increased when CIS is accompanied by at least one T2 lesion typical of demyelination on the baseline brain MRI (8, 36, 37). Another indicator is the identification of oligoclonal bands in the CSF; which doubles the risk of a progression to MS (38, 39). A CIS episode may be an acute or subacute demyelinating event that is unifocal or multifocal, and must have lasted for at least 24 hours (40). Typical presentations include: optic neuritis, brainstem syndromes that include isolated cranial nerve deficits, as well as partial motor or sensory deficits; and is usually followed by a complete or partial recovery (35, 40). As there are many conditions that may mimic the nature of these symptoms, a differential diagnosis must always be considered and excluded before a diagnosis of CIS can be made (40).

A diagnosis of MS can be categorized into clinical subtypes to describe the variety of disease courses that assists to provide appropriate treatment and prognosis (40). There are four MS subtypes that have been identified and defined: relapsing-remitting (RRMS), secondary-progressive (SPMS), primary-progressive (PPMS) and progressive-relapsing (PRMS). Majority of patients with MS (80-90%) present with subacute attacks to the CNS followed by a degree of remission to normal. This is followed by another randomized attack that usually occurs in a different CNS location, thus, presenting as RRMS (17, 34, 40). Recurrent attacks can vary in frequency and severity and may never revert to normal; often leaving the patient with a residual deficit. However, if there are residual effects, these deficits remain stable until the next exacerbation. The McDonald criteria indicate the requirement for the diagnosis of RRMS involves both clinical and radiological evidence or with evoked potentials to demonstrate dissemination in both time and space (29-31). Patients diagnosed with RRMS can transition to SPMS within an average of three decades after diagnosis, with risk factors associated with older age, longer disease duration, greater disability and greater number of relapses, whilst disease-modifying therapy exposure is associated with lower risk (41). Conversion to SPMS is

generally subtle and therefore often confirmed in retrospect and the conversion can involve a withdrawal of previous treatments due to a lack of effective disease-modifying therapies for SPMS(42). SPMS is characterised by at least one relapse followed by a progressive increase of neurological disability over time which can have plateaus of stability as well as punctuated by intermittent acute attacks in between. Typically, the MRI scans show fewer new lesion formations over time despite ongoing decline (40).

A less common form of MS, PPMS accounts for approximately 15% of patients (34). Patients diagnosed with this form do not experience acute exacerbations, but rather an insidious decline in neurological function after onset. For the diagnosis of MS there must be a minimum of 1 year's disease progression as well as a positive brain MRI, spinal cord MRI and positive CSF findings (17). The PRMS subtype describes a similar progression of disease function but may also be punctuated by relapses. Recently, it has been recommended that the PRMS term be made redundant as it is believed to overlap with other disease course subtypes i.e. a more aggressive form of RR (4, 40, 43).

1.2.4 Motor symptoms

The consequences of demyelination for saltatory conduction may explain many symptoms of MS. The symptomatic expression depends not only on the location of the MS lesions but also the type of conductions properties displayed by affected axons (44). Partially demyelinated axons induce: conduction impulses at reduced velocity, spontaneous discharge and ephaptic transmission (a trigger for many of the paroxysmal symptoms) (45). Lesions of the brainstem and cerebellar pathways produce precise clinical pathological correlations such as coordinated movement of eyes (46, 47). Affected areas in the spinal cord lead to alterations in sensory, motor and autonomic functions (48). However, most white matter abnormalities cannot always be linked to specific clinical symptoms (45).

The most common initial manifestation of MS demyelination is optic neuritis; 15-20% of patients initially experience it and about half of pwMS develop it at some point during the disease (17, 47, 49). Optic neuritis typically involves visual blurring or loss, eye pain and difficulties with colour contrasting that can evolve over hours to days and leave residual deficits (49). Other visual problems such as nystagmus and double vision are also common and occur as a result of demyelination in the brainstem, specifically cranial nerves III, IV and VI (17). In addition to visual problems, brainstem demyelination can produce a range of symptoms that include: facial weakness (cranial nerve VII), vertigo (cranial nerve VIII) or dysphagia, dysarthria, tongue weakness and swallowing difficulties (cranial nerves IX, X, XII) (17).

Motor symptoms related to MS typically involve focal weakness in the limbs and is often accompanied by other signs of motor neuron syndrome such as spasticity, hyperreflexia, stiffness, tremor, spasms, gait impairment and cramping (17, 50, 51). Numbness, tingling, itching and other forms of paresthesia and dysesthesia may also occur as initial symptoms in as many as 40% of patients (52).

1.2.5 Cognition

Cognitive deterioration is recognized as a prevalent and debilitating symptom of MS that is reported in up to 70% of all pwMS in early as well as late stages (53-55), and can have detrimental effects on emotional, social, employment status and yearly earnings to impact quality of life (54, 56). Although the heterogeneous nature of MS means that a variety of cognitive domains can be impaired, slowed information processing speed is the most prominent in MS (57). Attentional difficulties and executive functions have also been reported (58), whilst basic language abilities and general intelligence generally remain intact, even at more severe cases of the disease (59).

Information processing speed is the speed with which information is processed and integrated with other cognitive processes, to form a behavioural result (7). Deficits in this cognitive domain are prevalent within MS and are posited to be related to other cognitive domains, namely working memory (60). Slowing in information processing has been reported to be one of the first cognitive symptoms to be detected and especially affected in patients with a secondary progressive form (58).

Attention is a central cognitive process that is not a single mental function but involves many processes that facilitate the ability to focus on relevant information and inhibit irrelevant distractors (7). Complex aspects of attention such as selective, divided, and alternating attention are most often impaired in MS – whilst the simplest form, attention span, remains generally intact (59, 61). Attentional deficits are common (between 20-50% of patients) (58), and can occur even at the earliest stages of the disease to cause considerable debilitating effects on normal functioning (59).

Executive function refers to the ability to carry out a set of cognitive processes required for planning, goal-setting and adaptive behaviour. Abnormalities in such complex abstract thinking can be particularly debilitating for patients who are employed in high demanding jobs (62). Between 15-25% of patients display executive deficits, making executive abnormalities less frequent than the other domains previously described (63).

Once cognitive deficits are established, deficits tend to worsen with disease progression; and the degree of impairment is often associated with clinical phenotype; there are less cognitive impairments in relapsing-remitting than progressive forms (44, 64, 65). Cognitive deficits can also occur in the early stages of the disease and are likely to be subtle; reflecting the neural compensatory mechanisms that can effectively compensate for disruption to networks implicated in cognition (66-68). Often the identification of cognitive impairments in CIS predicts a greater chance of the conversion to definite MS (69). Although the mechanisms relevant to the development of cognitive deficits in MS are not completely understood, impairment involves a combination of processes that include inflammation, neurodegeneration, functional disconnection and cell atrophy (56). Cognitive abilities are usually assessed with a short neuropsychological battery as no single test alone can capture the widespread deficits that can occur (70).

1.2.6 Overview

MS is a challenging neurological disorder, with a largely unknown etiology that causes numerous heterogeneous symptoms and an unpredictable clinical course (71). Although MS remains an untreatable disease, newer immunotherapies yield long-lasting benefits and have significantly improved in efficacy over the last thirty years (71). Research into this debilitating disease continues with the aim to reduce disease burden and improve patient quality of life.

1.3 Hearing sensitivity and auditory flow in MS

The remainder of this review will focus on the various auditory deficits caused by MS-related demyelination and subsequent atrophy. Demyelination in MS can occur at any myelinated axon within the auditory processing pathway and several types of hearing impairments can result depending on which part/s of the auditory pathway are affected (72, 73).

1.3.1 Pure tone thresholds

Auditory processing begins with the pinna (outer ear) that funnels sound waves to the ear canal and middle ear. Mechanical sound energy waves transmit through the tympanic membrane, ossicular chain, and oval window resulting in a displacement of the basilar membrane and firing of cochlear hair cells. The neural discharge from the cochlear hair cells continues through the vestibulocochlear nerve tonotopically (a spatial arrangement based on sound frequency) and reaches the cochlear nuclei, located in the lower brainstem. The cochlear nuclei transfer the signal predominantly to the contralateral superior olivary complex and ipsilaterally to the lateral lemniscus. From there, the pathway connects to the inferior colliculus, thalamus and temporal lobe. In sensorineural hearing loss the damage occurs in the hair cells in the cochlear, the neural pathway, or both (74, 75). Cranial nerve abnormalities might affect up to 10-15% of patients with MS; amongst this, the vestibulocochlear nerve is rare (0.5%) (76).

To indicate if a potential lesion is located along the peripheral auditory pathway, behavioral audiometric evaluation commonly consists of pure tone air testing. It can help rule out lesions along the auditory nerve and illustrate the need to test for a more centralised auditory processing disorder (77). To evaluate hearing loss in MS, pure-tone thresholds are measured at the frequencies 0.25, 0.5, 1.0, 2.0, 4.0 and 8.0 kHz in each ear (10).

There is no basic agreement in the literature regarding the prevalence and nature of audiometric hearing loss associated with MS (9, 78). After reviewing multiple studies, Noffsinger et al., (1972) reported that prevalence is extremely variable, ranging anywhere from 1 to 86 percent (79). In such MS-related cases, hearing loss has been described as unilateral and bilateral, sudden and progressive, symmetrical and asymmetrical, mild and severe, acute and chronic (9, 80). Such variety in prevalence and descriptions of hearing loss are likely to reflect the heterogeneous

nature of MS and the array of experimental methods used (9, 78, 81). Acute influences of MS on pure-tone thresholds are well documented as a feature of the initial attack or exacerbation of symptoms (82, 83). However, most of the inconsistencies arise in chronic hearing loss studies (81). Chronic cohort studies have illustrated a variable pattern of affected frequencies on audiometry describing: high frequency, "dome shaped" and low frequency configurations (9). Despite variability, the general view is that when pwMS are not experiencing an acute exacerbation, difficulty with hearing is a rare complaint (10).

1.3.2 Acute influences on pure tone thresholds

Acute hearing impairment in pwMS is relatively rare. In fact, it is difficult to ascertain whether the prevalence of acute hearing loss in an MS cohort exceeds that of the general population (9). Case reports and cohort studies investigating acute hearing loss in pwMS describe it to be typically unilateral, mild in nature (i.e. ≤ 40 dB) with a good prognosis (11, 78). Within such hearing loss cases, some patients fulfill the criteria for sudden onset sensorineural hearing loss (SSHL), which is defined as a hearing loss of at least 30 dB in three sequential frequencies in the standard pure tone audiogram occurring within 3 days or less (76, 84).

Several large cohort studies have described hearing loss in pwMS during the midst of an acute exacerbation of CNS symptoms. Fischer et al. (1985) reported 12 out of 705 pwMS (1.7%) presented with hearing loss during a relapse of the demyelinating disease. Hearing loss was found to be unilateral in all 12 cases but one, and chronically persisted in only one patient. Similar trends were mirrored in another series published almost 30 years later (83). Leite et al. (2014) identified 7 out of 405 pwMS (1.7%) had acute hearing loss related to MS outbreaks. Unilateral involvement was also a feature described in majority of cases (5 out of 7); two patients were identified with hearing loss during the initial manifestation of MS; and remaining participants experienced deafness up to 19 years after disease onset (85). Other larger series note higher instances of acute hearing loss presenting early in the course of the disease. Hellman et al. (2011) retrospectively evaluated 253 patients over a 6-year period and 11 (4.35%) had sudden hearing loss early in the course of the disease (<5 years). Seven patients had hearing loss as the presenting symptom and most hearing loss cases resolved with a residual deficit in only two cases; suggesting that MS induced hearing loss has a good prognosis (76). Although the reported prevalence of acute deafness in pwMS varies amongst series, most studies claim it to be of low prevalence (< 10%) (86, 87). Refer to **Table 1.1**.

Table 1.1. Summary of studies investigating acute influences of MS on hearing loss		
Investigators	No. with acute hearing loss/ total	Features of hearing loss
Fischer et al.(1985)	12/705	Unilateral in all cases but 1. Chronic in 1.
De Seze et al. (2001)	14/400	In all but 2, deafness associated with disease exacerbations. Chronic in 1.
Hellman et al. (2011)	11/253	Initial symptom in 7, Residual deficits in 2.
Leite et al. (2014)	7/405	Unilateral in 5. Initial symptom in 2.

Table 1.1. Summary of investigations into the acute influences of MS on hearing loss. Several large cohort studies describe hearing loss in pwMS during an acute exacerbation of CNS symptoms. Despite varying reports, all studies report a low prevalence (<10%) of hearing loss that is generally unilateral and transient.

In addition to cohort studies, isolated case reports have appeared in the literature since 1888 (80) and have shed light on the relationship of pure-tone behavioral responses to objective measures such as MRI or ABR. Investigators commonly ascribe the presence of inflammatory demyelinative lesions of the eighth cranial nerve as the cause of acute hearing loss (82, 83, 88). In fact, one of the first acute hearing loss cases in an MS patient verified with an MRI showed lesions in the 8th nerve root-entry zone and cochlear nucleus (89). However, not all cases of hearing loss indicate a lesion located at the lower levels of the pathway. A rare case of cortical deafness in MS was documented in a patient who exhibited bilateral temporal lobe lesions with no lesions in the lower level of the auditory pathway (90).

Other investigators determining the site of lesion responsible of MS hearing loss have used measurements of the auditory brainstem response (ABR). Such studies describe abnormal evoked potentials that implicate the involvement of the vestibulocochlear nerve close to the pontomedullary junction (9). However, it is difficult to ascertain whether such abnormal ABRs predate acute hearing losses. Shea and Brackmann (1987) reported profound unilateral hearing loss as the initial symptom of a 20-year-old woman, which ultimately led to a diagnosis of MS. The patient's hearing had returned to normal after 3 months but the ABR remained abnormal in that ear (91). Similar reports of abnormal ABRs despite the presence of normal peripheral sensitivity further illustrate that the relationship between lesion location and electrophysical findings to MS-induced hearing loss remains complex and unclear.

1.3.3 Chronic influences on pure tone thresholds

There is less consensus about chronic influences of MS on hearing loss compared to acute influences (9, 81). Investigations in chronic hearing impairment in the MS population face major methodological challenges which have produced extremely inconsistent findings (refer to **Table 1.2**). To describe hearing loss in the MS population, it is imperative to consider the potential confounding factors of age, sex and noise exposure on the audiometric status of the general population (92). PwMS within an older age bracket are likely to display some degree of hearing impairment

due to presbycusis (93). Furthermore, age-related hearing loss is also gender specific: males are generally characterized by high-frequency loss and women tend to have lower frequency hearing loss (92, 94).

In reports that draw comparisons to the normal population, the studies are ambivalent to whether MS chronically affects pure tone thresholds or not. In the earliest study of this kind, the thresholds of 78 pwMS were studied and elevated thresholds occurred more often at lower frequencies (68% of pwMS) (95). However it has been pointed out that although participants were age matched to controls, they were not sex matched (9, 81). Lewis et al. (2010) reported that MS chronically affects pure-tone thresholds at *both* high frequencies and low frequencies (0.25, 0.5 and 0.75 kHz) in 47 pwMS. But once again, the MS cohort was not one-to-one gender matched with controls and considerably more war veterans were included in the MS cohort which could bias the results as veterans between 48-59 years have slightly, yet significantly, higher pure-tone thresholds (2 to 3 dB) at higher frequencies than non-veterans (96).

A well-controlled report by Doty et al., (2012) established that MS is not chronically associated with pure-tone hearing loss. In the largest group study of its kind, 73 pwMS who had the disease for over 7 years were individually matched based on age, sex, and ethnic background. No statistical differences in pure-tone thresholds of pwMS and controls were found. As Doty et al. (2012) pointed out, their findings are in accordance with other case-control studies conducted by Cohen and Rudge (1984) and Coelho et al., (2007), but contrast with others (Dayal & Swisher, 1967; Lewis et al. 2010; Simpkins 1961). The results of this study affirm the importance of controlling for basic variables when making inferences about hearing loss in the MS population. It is also clear that hearing impairments in pwMS are a rare complaint, and if they do occur, they are likely to be acutely influenced by a CNS exacerbation rather than chronic influences.

Table 1.2. Summary of studies investigating chronic influences of MS on hearing loss

Investigators	Number of participants	Sex (Male/Female)	Age range	Hearing loss configuration
Simpkins (1961)	78	29/49	19-83	Rising or dome-shaped
Dayal & Swisher (1967)	22	9/13	32-66	High frequency or flat
Cohen & Rudge (1984)	44	16/28	15-50	Rising
Jerger et al. (1990)	62	21/41	21-61	Dome-shaped
Lewis et al. (2010)	47	26/21	35-65	High and low frequencies
Doty et al. (2012)	73	21/52	Mean = 45	No configuration

Table 1.2. Summary of investigations into the chronic influences of MS on hearing loss. Several cohort studies report contradicting audiometric hearing loss profiles in the MS population. A well controlled study by Doty et al. (2012) suggests that there is likely to be no effect of MS on pure tone audiogram configuration.

1.4 Central auditory processing deficits

Despite having normal pure-tone thresholds, studies have shown that pwMS often display abnormal electrophysical responses to auditory stimuli and perform abnormally in several auditory tasks (10). In fact, Musiek et al. (1989), reported that 33% of patients with normal pure-tone thresholds subjectively reported of hearing difficulties (97). A lack of correlation between subjective reports and normal pure-tone thresholds, in parallel with abnormal AEPs, was hypothesized to reflect central auditory processing disorders (CAPD) in people with MS (97).

CAPD is defined as dysfunction of the central auditory nervous system that can manifest as perceptual auditory processing problems which are not due to peripheral hearing problems (98). CAPD has several suspected causes, including neuromaturational delay, neurological insult of the central auditory nervous system, and central presbycusis (99). As MS can cause neurological insult at any level of the auditory pathway (**Figure 1.1**), the final auditory percept is likely to be affected. It should be noted that the brain is not organised as a straight-forward hierarchical system in which information is processed sequentially in ascending order of the CNS, rather, it is a distributed network whereby parallel processing allows higher order factors such as attention and memory to influence bottom-up factors. In fact, primary complaints of CAPD may be in the auditory domain, but functional deficits can manifest in areas related to attention, language, communication, and learning (100). Supporting this notion, many of the symptoms of CAPD are presented in many other disorders such as autism spectrum disorder (ASD), dyslexia and attention deficit disorders, with evidence indicating that reduced higher-order cognitive abilities may play a role in CAPD (99). Furthermore, Tomlin et al., (2015) demonstrated that many children who have poor CAPD test performance also have lower scores on cognitive tests.

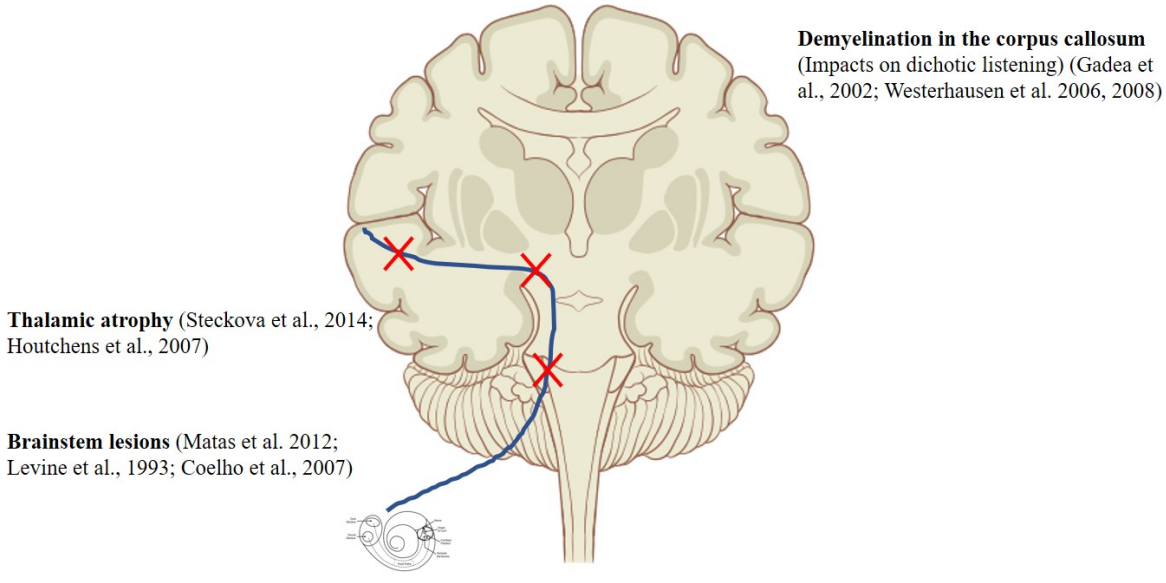


Figure 1.1. Neurological insult caused by multiple sclerosis pathology at any point in the central auditory pathway can manifest as perceptual auditory processing problems.

Estimations of the prevalence of CAPD in the general population vary widely between 0.5 – 10%, and approximately 3-5% in children (101). Such a wide prevalence rate is likely to reflect the difficulty in the diagnosis of CAPD (101) which requires a multidisciplinary approach between audiologists, speech pathologists, psychologists and neuroscientists employing both electrophysiological and behavioural assessments (102). Four main auditory processes are measured in the diagnostic CAPD battery, and these include (1) binaural interaction, in which the two ears receive complimentary input which is integrated to support the perception of spatial localisation; (2) temporal processing, which is the processing of temporal aspects of sound; (3) perception of monaural speech degraded through filtering; and (4) dichotic processing, in which speech is simultaneously presented to each ear (100).

Complicating diagnosis, the criteria for CAPD is not universally accepted, and confounding effects of non-auditory factors such as cognitive ability, impact on CAPD test performance. One solution to mitigate this confounding effect many be to evaluate patients with a comprehensive neuropsychological assessment battery, however the approach to evaluating CAPD remains an evolving process of refinement.

1.5 Auditory flow in MS: Auditory evoked potentials

Auditory evoked potentials (AEPs) are physiological measures which reflect the neuroelectric activity within the central auditory pathway, from the auditory nerve to the cerebral cortex, in response to an acoustic stimulus (103). Such responses can provide more subtle information on the integrity of the afferent auditory pathways than pure tone audiometry (104, 105). Electrodes positioned on the scalp record responses to the sounds and are observed as a reading on an electroencephalogram (EEG). The most studied AEPs are the brainstem auditory evoked potential (BAEP), the auditory middle latency response (AMLR) and the cognitive potential (P300)(106). The BAEP consists of seven waves and assesses the integrity of the auditory pathway from the auditory nerve to the brainstem (106). The AMLR, generated by the activation of several subcortical structures, assesses the functional integrity of the auditory pathway above the level of the brainstem in cases with suspected lesions (107). The cognitive potential (P300) is an endogenous potential elicited in the process of higher level processes such as attention, auditory discrimination, memory and decision making (108, 109). Response analysis criteria generally evaluate the latency (milliseconds (ms)) and amplitude (microvolts (μv)) values of the wave components. Several authors have highlighted the importance of using BAEPs in the MS population due to its ability to detect brainstem dysfunction that may evade detection morphologically in MRI scans (the so called clinico-radiological paradox) (110). Furthermore, evoked potentials (EP) have proven clinically useful as they reliably predict disability in pwMS; specifically, the index of global EP alteration which combines alterations in visual, auditory, motor and somatosensory evoked potentials has significantly correlated with EDSS scores at 1, 3 and 5 years after follow up (111).

1.5.1 Brainstem auditory evoked potentials

In a large study involving 202 pwMS, Chiappa et al. (1980) reported 64 patients (32%) had abnormal BAEPs in response to monaural click stimuli presented at 10 clicks per second (/s). The most common BAEP abnormality in their MS population was the absence or abnormally low amplitude of wave V (87% of BAEP-abnormal pwMS) and increased inter-wave latency between waves III-V (28% of BAEP abnormal pwMS) (110). Several studies investigating BAEPs also reported substantial alterations in wave component V, both with respect to decreased amplitude and increased latency (104, 112). Matas et al. (2010) reported a higher prevalence of abnormal BAEPs in their clinical population with 68% of patients displaying abnormal results compared to controls (106). Similarly, an earlier study conducted by Kofler et al. (1984) also reported a high incidence of 68% in their MS subgroup (113). Variances in abnormal BAEP incidences between studies may reflect differences in click stimuli being used; Matas et al (2010) monoaurally presented click stimuli at a rate of 19.9/s, which is faster than Chiappa et al.'s (1980) 10 clicks/s stimulus. Stockard et al (1977) and Robinson and Rudge (1977) had reported that faster rates of stimulation revealed a higher incidence of abnormalities (112). Despite variances in incidence, abnormalities similarly involved increased latencies of waves III and V, increased I-III, III-V and I-V interpeaks and the absence of one or more components (106, 110, 114). Thus, the most commonly observed trend in abnormalities was the selective loss of late waves with preservation of early waves. Several authors concluded that majority of the conduction abnormalities are likely because of the effects of demyelination on neural timing (110).

1.5.2 Auditory middle latency response

The AMLR consists of a set of positive "p" waves and negative "n" waves (115). The first negative wave is referred to as Na, followed by the positive wave, Pa. The Na component is believed to originate subcortically from the midbrain, thalamus or thalamocortical radiations; whilst the Pa component related to activation of both supratemporal auditory cortices (115). In a quantitative analysis of 25 pwMS, AMLRs were elicited by a click stimulus presented monoaurally at 70 dBHL, at a rate of 9.9 clicks/s (106). No statistically significant differences were observed between groups for the latency of wave Pa or Na-Pa amplitude. There was however, a significantly abnormal Na latency with 92% of pwMS having an abnormal delay (n=23). Other investigations also reported abnormalities of the AMLR response elicited by pwMS; with the occurrence of abnormalities in 73% (107) and 45% of pwMS producing abnormalities of latency of one or more AMLR components (112). Variances in such investigations to the prevalence of AEP abnormalities can be traced to variation in criteria of abnormality (116). Some investigators employ the criteria of ±2; ±2.5 or ±3 SD of the normative data for defining the latencies as abnormal (107, 116).

1.5.3 Cognitive potential (P300)

The P300 wave is commonly elicited by a 'odd-ball' discrimination task which consists of a series of frequent and target stimuli (auditory or visual) randomly administered in the proportion of 4:1, respectively (108, 117). The participant must engage attention and memory during the performance to indicate when they can identify the target stimulus. Hence, the P300 wave is associated with anticipation of the stimulus, decision making and control of

behavior (108). Latency and amplitude describing wave P300 are used as neuropsychological indicators of cognitive impairment and these measurements have been reported in pwMS to assess the diverse cognitive profile in different forms of the disease (RRMS, PPMS, SPMS) (106, 108, 109). Majority of investigations report that pwMS had abnormal findings in at least one parameter that describes the P300: the amplitude and/or latency of the P300 wave (106, 108, 109). Specifically, Magnano et al. (2006) described P300 latencies significantly increased (>2 S.D) and also lower P300 amplitudes with respect to age matched normative data (109). The presence of abnormal P300 waves suggest that pwMS have deficits in cognitive domains such as memory, attention and auditory discrimination due to impairment in cortical regions of the auditory pathway. The latency marker is an indicator of prolonged information processing time whilst reduced amplitude may reflect disruption in the activity of cortical regions (108).

1.6 Processing cues to auditory space

As MS is a CNS disorder, it is likely that CNS functions are more vulnerable than pure-tone thresholds in pwMS. Binaural hearing is a CNS function that integrates sounds from both ears to make a judgement about the location of a detected sound in direction and distance (118). The two ears receive slightly different information generated by the interactions of sound waves with the external physical shape and location of the ears, head and body (118). Along the frontal azimuth plane (the horizontal dimension), two acoustic cues are dominant for localization: i) interaural time differences (ITDs) and ii) interaural level differences (ILDs) (119). In practical terms, any source displaced from the sagittal plane results in the ear closest to it receiving the sound slightly quicker and louder than the other ear. Investigators have used lateralisation tests and assessed "just noticeable differences" (JNDs) for ITDs and ILDs to assess a MS participant's ability to locate stimuli; commonly trains of clicks, noise bursts with difference frequency content, or tones.

1.6.1 Lateralisation tests

The lateralisation test is one test employed to assess binaural function for localising sound in space and it has elicited patterns of abnormality in position judgement for the MS population (10, 120-125). Since lateralisation depends on the integrity on the brainstem, it is logical to expect that the MS population would experience difficulties with lateralising sound (9). Various methods have been employed to test an MS participant's ability to indicate where in their head they perceive the location of a binaural stimulus; usually a train of dichotic rarefaction clicks delivered to earphones with various ITD or ILDs. One method requires the participant to report the position of each click train on an evenly spaced numbered scale used to parallel the listener's perceived auditory space. With the scale as the basis for judgement, participants can use any integer from 1 to 9 to report sound position (number 1 would signify complete lateralisation of the sound to the left ear and 9 to the right). Another method requires the participant to match a pair of rarefaction click trains. The first train of clicks with a fixed ITD serves as a reference and the participant's task is to match the perceived position of the second train to the first by rotating a knob that changes the ILD parameter of the

second train. Trials are repeated with the participant adjusting the ITD to match the position of the reference with a fixed ILD.

In a pilot study (123), and an expansion on this investigation (125), Furst et al. (1995; 2000) classified the lateralisation performance by pwMS into three major categories. *Group I* comprised of MS listeners whose performance was within normal limits (within 2 SD from the control mean) for both ITD and ILDs. *Group II* described patients whose performance with ITD stimuli was poorer (more than 2.5 SD from mean), but with normal ILD detection. Group III consisted of patients who abnormally lateralised dichotic clicks using both ITDs or ILDs. Abnormal perception of all stimuli (time and/or level) were further categorized into two types 1) centre-orientated or 2) side-orientated (123-125). A centre-orientated lateralisation was obtained by patients who were biased towards centre positions. These participants tended to incorrectly perceive the sounds located at the side as being midline to their head. In contrast, those participants with a side-orientated abnormality tended to perceive dichotic stimuli as being located to the sides of their heads and very rarely in the centre. Similar centre and side orientated behaviours during lateralisation were also detected in stroke patients with ischemic lesions (124, 125).

1.6.2 Relationship between lateralisation performance and MRI lesions

Efforts have been made to describe the relationship between MS lesion site and abnormal lateralisation. Furst et al. (1995, 2000) were amongst the first studies to implicate the pontine auditory pathway in sound lateralisation. Patients who performed normally in the position judgement experiments had no detectable lesions involving the pontine auditory system. When lesions were restricted to the caudal pons, psychoacoustic performances consisted of normal ILD but abnormal ITD position judgements; and click lateralisation for ITD was always centre-orientated. For lesions rostral to the trapezoid body, lateralisation was side-orientated for both ITD and ILD performance. For lesion sites involving both the caudal pons and more rostral auditory structures, lateralisation testing was a centre orientation (124).

1.6.3 Interaural discrimination tests

Another way to evaluate binaural functioning is to determine the 'just noticeable difference (JND) or smallest interaural difference required to make an accurate judgement about sound location (10). Levine et al. (1993) estimated time and level JNDs of 38 pwMS by presenting two successive noise bursts that were either high-frequency (high-pass filtered noise bursts (>4000 Hz)) or low frequency (low-pass filtered noise bursts (<1000 Hz)). The first stimulus provided a reference intended to be perceived in the midline (both ears receive identical sounds) and the following dichotic stimulus was perceived to one side of the reference burst depending on the ITD or ILD cue. The participant's task is to indicate in which direction the two successive stimuli move relative to each other. Throughout testing, trials become easier (larger interaural differences) or more difficult (smaller interaural difference) depending on the

accuracy of responses. The JND was defined as the mean of the 6 'turnarounds', i.e. when the test stimulus had been changed from difficult to easy and vice versa (122, 126).

Comparisons of JND values to control participants indicated that pwMS obtained abnormally large JNDs for both ITD and ILD cues. The most sensitive indicator of auditory dysfunction was found to be the ITD cues; more specifically, the ITD cues for high-frequency sounds. In fact, ILDs appeared to be largely left intact by the disease process. 70% of all pwMS had abnormal JNDs using ITD for high frequency sounds compared to 40% for the low-frequency ITDs. Only 11% of participants had abnormal low frequency ILDs and 8% for high frequency ILDs (126). Levine's (1994) results mirror this study with a reported 42 out of 76 (55%) abnormal JNDs required to detect ITDs and only 7 out of 76 (9%) abnormal JNDs for ILD stimuli (127). Aharonston's 1998 results also found that high noise band JND for ITD was always abnormal whenever any other JND was abnormal in an MS participant (124). The discrimination of ITDs by pwMS was a comparatively major dysfunction compared to ILD cues. It was concluded from the results that most mechanisms involving precise neural timing were affected by the disease whilst other auditory mechanisms that don't (e.g. interaural level discrimination and pure tone thresholds) are largely left intact by the disease process. In fact, the auditory system is probably the most sensitive neural system to temporal features of stimuli, with some aspects requiring resolution in the microsecond range (127).

When investigating the correlation between JND measurements and lateralisation performance in a group of 9 pwMS, Aharonston et al's (1998) results indicated that abnormality in lateralisation did not necessarily indicate an abnormal JND. Their results indicated that if any JND or lateralisation task was found to be abnormal, then the lateralisation for high noise band with ITD will always be abnormal. They concluded from their results that lateralisation tasks with high-frequency stimuli may be the most sensitive detectors of abnormality than JNDs (124).

1.6.4 Relationship between interaural discrimination and brainstem auditory responses

Efforts have been made by investigators to characterize the relationship between psychoacoustic behavioural data and objective measures. In a combined psychophysical-electrophysiological study, Hausler and Levine (1980) tested the ability of pwMS to discriminate interaural time and level differences of binaural noise bursts and measured their AEPs in response to a train of 10/s monaural rarefaction clicks at 70 dB HL. All pwMS who performed poorly in the ITD discrimination test also had abnormal BAEPs in at least one ear. In contrast, no obvious correlation was found between AEPs and the ability to discriminate ILDs. Additionally, several of the patients with poor ITDs presented with a lateralisation bias and consistently lateralised the stimulus towards one side; usually to the side of the ear that produced the better AEP (128). Confirming such findings, a later study conducted by Van der Poel et al. (1988) described a similar relationship and demonstrated a correlation with specific components of the waveform. Specifically, abnormalities involving the absence or prolonged latency of wave III of either monaural BAEP was correlated to poor ITD discrimination (129). Like Hauler and Levine (1980), no evidence was found of any association between interaural level discrimination defects and the abnormalities of monoaural BAEPs.

MRI lesions involving the auditory pathway have been correlated to abnormal BAEPs (127). Brainstem lesions were detected in 5 of 16 participants. One participant had a pontine lesion, but this lesion did not overlap with the auditory system. In the 4 participants involving lesions in the auditory brainstem pathway, all of them had abnormal BAEPs to the point of having unidentifiable waves (127).

1.7 Temporal processing in MS

It's clear that pwMS have trouble using timing cues to locate sound in space- but what about other functions that use the temporal domain? Temporal processing refers to the accurate encoding of durational characteristics within complex acoustic signals such as speech (86). Deficits in temporal processing may lead to a distortion of rapidly presented spectral cues that naturally occur in speech with background noise and sound localisation. MS abnormalities in temporal processing have been reported at varying levels of the auditory pathway; from the time sensitivity of the first-order neurons in cochlear nuclei within the brainstem (i.e interaural discrimination tasks) (125, 126) to the processing of complex auditory information at the cortical level (i.e. speech-in-noise paradigms) (86, 130).

1.7.1 Masking level differences

The masking level difference (MLD) is a psychoacoustic phenomenon in which the ability to detect a signal in noise is improved when the phase of either the signal or noise is reversed by 180º (77). The degree to which the out-of-phase threshold is superior to the in-phase threshold is referred to as the MLD (131). Several studies have employed the use of MLDs as a method of assessing the integration of binaural, temporal information; and it is particularly sensitive to functioning at the lower levels of the brainstem. MLDs can be measured monoaurally and binaurally, and tests are typically conducted with tones at low frequencies and gaussian white noise as the masker (79, 97, 127, 132).

In the earliest work done on MLDs in the MS population, Noffsinger et al. (1972) utilized both 500 Hz tones and spondaic (two syllable) words as the test stimuli. The MLDs were obtained by comparing the amount of masking produced under two conditions: 1) both the noise and test stimulus in phase with itself at the two ears (S0N0); 2) the noise in phase with itself but the test stimulus 180º out of phase with itself at the two ears. 23 participants (almost half) had abnormal MLDs of 7 dB or less when detecting the 500 Hz tone in the masker (narrow band noise, 80 dB SPL overall level). In comparison, 95% of the control population had MLDs of 8 dB or more; suggesting that MS listeners were not able to gain as much of an advantage in detecting the sound stimulus when there were differences in timing of the stimulus being presented. Furthermore, in the spondee MLD test, 30/42 pwMS achieved abnormally small MLDs of 5dB or less (95% of the control population obtained MLDs greater than 5) (79). A relatively recent study by Lewis et al. (2012) reported abnormal MLDs in 28% of pwMS, and identified MLDs as a promising central auditory processing test for screening pwMS. It should be noted however, that it was a preliminary investigation and authors recommended that references standards needed to be established in a much larger sample population before

use in the clinic (12). Other studies that also tested MLDs found variable numbers of pwMS who performed abnormally, ranging from reports of 40% (132), 42% (105), 45% (116), 50% (97), through to 58% (133) of pwMS obtaining smaller MLDs than the controls (for a summary, see **Table 1.3**). When the detection rate of abnormalities produced by the MLD test was compared to several other common audiological tests, MLDs proved to be amongst the most sensitive and useful measures for this clinical population (12, 97).

Table 1.3 Summary of Studies Evaluating the Masking Level Difference in MS

Investigators	Number of participants	Normal Criterion (dB)	Abnormal results (%)
Noffsinger et al. (1972)	61	>7	49
Olsen et al. (1976)	100	>7	44
Matathias et al. (1985)	43	>8	42
Musiek et al. (1989)	33	>7	50
Hendler et al. (1990)	15	>6	40
Jerger et al. (1990)	62	>7	45
Levine et al. (1994)	37	>4	11
Lewis et al. (2012)	26	>10	28

Table 1.3 Summary of investigations into the effects of MS on masking level difference (MLD). Several studies have assessed MLD performance in pwMS and have demonstrated that it is a sensitive measure of temporal processing. Reports range from 11-49%, with some variability attributed to differences in criterion for abnormality.

Mustillo (1984) postulated that a contributing factor to abnormal MLD performance could be a deficit in temporal processing related to delays in signal transmission within demyelinated auditory pathways (11). Nerve impulses may still propagate, but the temporal relationship between ears in neural conduction leads to a deficient processing of differences in phase. In this regard, it is logical that Levine et al. (1994) found a high correlation between MLD performance and the discrimination of ITDs at high frequencies. Out of the 37 participants tested, 4 of them (11%) had abnormally smaller MLDs than controls. All 4 of them also had abnormal JNDs for ITD. The correlation between electrophysiological responses and MLDs in pwMS is less clear (10). Hendler et al. (1990) reported that participants with abnormal MLDs were more likely to have bilateral abnormalities in both BAEPs and AMLRs (132). Levine et al. (1993) found less correlation between the two measures since there were patients whose BAEP was abnormal but with normal MLDs. Furthermore, there were some patients who performed abnormally in the BMLD task who had normal BAEPs (126).

1.7.2 Frequency pattern test

Another approach to assessing temporal resolution is to use the frequency pattern test (FPT). The FPT is composed of sequences of three tone bursts presented to each ear. The sequence comprises of a high frequency and a lower

frequency tone and the listener is asked to describe the pattern by using the words 'high' and 'low' (e.g. 'high-high low'). Frequency discrimination would begin in the cochlear on the basilar membrane and detection would continue in the auditory nuclei in the brainstem. The conscious recognition of frequency differences would occur in the primary cortex and the stimulus patterns would be routed to auditory association areas for further processing (77, 134). With such widespread auditory activity involved in frequency pattern processing, it's logical to expect that participants with CNS involvement, specifically lesions in the cochlear, brainstem and cerebrum, performed the test abnormally. 45% of participants with brainstem involvement (n=22) and 83% of participants with cerebral lesions (n=29) were considered to have abnormal results (a score below 75% for either ear) (135). Lewis et al. (2012) also conducted a FPT but unlike Mustek (1987), concluded that it wasn't a very sensitive measure for detecting MS auditory deficits as no significant differences were reported in performance between controls and pwMS (n=26) (12). It is possible that the patients tested in this study did not have lesions that infringed on any of the brain areas vital to the central processing or transmission of the sequence and response (12). Additionally, compensatory cortical activity in pwMS has been heavily reported in the literature, and might be why the FPT did not detect FPT abnormalities in this MS patient sample (136, 137). Mustek et al. (1987) concluded that no one central auditory test is sufficient for detection of all lesions involved in the auditory pathway and highly recommended that the use of a central auditory test battery is always used (135).

1.7.3 Gaps in noise test

The gaps-in-noise (GIN) test is a common procedure used as a means of measuring temporal resolution, shown by Musiek and colleagues to be sensitive to patients with central auditory processing involvement (138). The monoaural test is composed of a series of broadband noise segments of 6 seconds in duration. Each segment of noise contains 0-3 silent gaps which vary in duration (2-20 msec) and patients are required to push a response button if they perceive a gap separating two successive signals. Measures used to determine performance on the GIN test are the approximate gap detection threshold and the overall percent correct (77).

GIN and other gap detection paradigms have been studied in patients with MS. In 1990, Hendler et al. (1990) reported that 2 of 15 clinical participants had significantly elevated gap-detection thresholds when gender, age and hearing status were considered. They concluded that of all the measures tested (MLDs, ABRs and MLRs) gap-detection performance was least affected by demyelinating lesions (132). Lewis et al. (2012) also noted that the GIN test was the least sensitive measure in their test battery (12). Low levels of abnormality in gap detection thresholds were reported in other MS populations: 4 out of 10 participants (86); 9% in left ear and 11% in right ear (12). Rappaport et al. (1994) reported that of the four participants with elevated gap-detection thresholds, three of them had lesions restricted to left forebrain auditory pathway sites. This contrasts Hendler et al.'s (1990) findings that monaural temporal acuity is disrupted by pervasive lesions located both in the upper brainstem and auditory cortex. The seemingly robust resistance to neural disruption by the gap-detection task was speculated to be due to compensatory mechanisms involving parallel processing and alternative routes of transmission (86, 132). As gap-detection relies on

accurate perception of minute temporal cues, the task has been thought to be related to speech perception (11); therefore, it is not surprising to expect that such an essential function would have alternative pathways for input to reach cortical structures (132).

1.8 Speech perception in MS

Speech perception has received comparatively less attention in MS than aforementioned psychophysical tasks, despite it being functionally more critical in the everyday world of human beings. A common audiologic test is the speech audiometry test that measures a patient's word recognition abilities. Speech recognition thresholds (SRTs) are defined as the softest level in dB HL a patient can hear and correctly repeat two-syllable words 50% of the time. Generally, pwMS perform normally in this task when the standard clinical level (70dB above SRT) is used (10). However, verbal messages in today's world often occur in the presence of natural, but unfavourable background noise and listeners must be able to develop sensory, cognitive and neural resources for handling noise to achieve successful communication (139). Normal hearing individuals take advantage of spectral and temporal "dips" for speech intelligibility and discrimination from background noise (140). Other clinical populations with known and probable deficits in the CNS, such as Autism Spectrum Disorder (ASD), have shown difficulties using these dips (141). It is therefore reasonable to postulate that pwMS with confirmed demyelination in the CNS may result in anomalies in understanding speech in background noise. Some studies have suggested that a high percentage (33-69%) of individuals with MS have trouble understanding speech when listening is accompanied with a degree of background noise (130, 142, 143). However, many of these studies did not make comparisons with normative data and were conducted monoaurally. As a general rule, tests sensitive to central auditory processing functioning such as speech in background noise and dichotic speech are often abnormal in pwMS; whereas measures of speech understanding in a quiet environment are more likely to be normal.

1.8.1 Speech in silence

Understanding single syllable phonemically balanced words is generally normal in pwMS. LeZak and Selhub (1966) tested 30 patients with MS and reported mean discrimination scores of 94 and 95 percent in the left and right ears, respectively (144). Other investigations also observed normal discrimination performance (145, 146). In instances where studies have reported abnormal phonemically balanced word scores, the varying percentages are relatively small, ranging from 3% (146) to 7% (79).

A partial explanation of why some patients experience degraded speech comprehension despite normal pure-tone audiometric profiles could be due to difficulties in discriminating changes in pitch. Quine et al. (1984) investigated the ability of participants to discriminate between computer-generated speech-like sounds consisting of three formant frequencies; approximated to mimic human speech. Participants were required to use pitch cues to identify the target sound that differed from the other two identical sounds. 9 of 25 patients had significantly poorer

discrimination performance than controls (147). This is consistent with other findings of selective impairment of sensitivity to changes in frequency in pwMS despite normal pure-tone thresholds (148).

1.8.2 Dichotic speech paradigms

In a test of dichotic listening (DL) the participant is presented with two different auditory stimuli to the left and right ears. Depending on the task, the participant may have to a) integrate binaural auditory input and repeat what was presented in both ears or b) segregate binaural input and ignore information in one ear whilst repeat what was said in the other. The ability to integrate and segregate such information is crucial for accurate speech perception in difficult listening situations (77). A major component of this test requires transmission of information between the two cerebral hemispheres via the corpus callosum (CC) (149). In patients with MS, lesions characteristically cluster around the ventricles and periventricular white matter. As the CC forms the roof of the lateral ventricles, MS lesions will often be located in and around the CC (150). These lesions are well situated to raise the possibility of there being an association between MS-related CC atrophy and functional consequences requiring interhemispheric interaction, such as dichotic listening (DL) (151-154).

The most commonly employed dichotic speech paradigm is the consonant-vowel (CV) Nonsense Syllable test. A pair of consonants are presented to the listener and they must repeat back the two vowels that they can hear. In a healthy control, it is usual to expect a superior report for verbal stimuli presented to the right ear compared to the left ear. This right-ear advantage (REA) can be explained by both anatomical and attentional properties of the auditory system (149, 155). However, in pwMS, pathological DL performances have been found in the form of a pronounced laterality effect compared to controls (145, 151, 153, 156-159). The earliest study in 1983 conducted by Jacobson Deppe and Murray reported reduced left ear scores in 20 pwMS in comparison to normative data. The mean percentage correct scores for the right and left ears of pwMS were 76.5% and 45.8% respectively whilst controls scored 71.8% and 58.3%. Although no statistical difference could be found between the right ears of the two groups, left ear scores were significantly different (145). Rubens' et al. (1985) results confirmed the findings of Jacobson and associates of a left-ear suppression on a nonsense C-V dichotic listening test in pwMS (156). In addition, both studies highlighted how only some dichotic tests, such as the C-V test, were sensitive to MS whilst others were not.

Two other tests employed by Jacobson et al. (1983) were the Staggered Spondaic Words (SSW) test and the Synthetic Sentence Identification (SSI) test and they were considerably less sensitive to MS because of very high-performance left-ear scores of both treatment groups. Performance was abnormal in 22% (97) and 10% of pwMS (145). Rubens et al. (1985) concluded that the greater the spectral/temporal overlap of competing dichotic pairs, the more sensitivity to the presence of MS. Minimally contrasting words widen the difference in performance between controls and pwMS by minimizing floor and ceiling effects and may provide the best type of stimulus material for a clinically useful test (156).

In addition to left-ear suppression in MS (145, 154), other dichotic studies found that pwMS were superior to controls in right ear scores (158), whilst others reported both effects (157). In some studies, enhanced right ear performance has been attributed to release from interference from left-sided input (158). Such findings are consistent with the hypothesis that functional callosal disconnection occurs in MS.

1.8.3 Speech in noise paradigms

Individuals with MS are susceptible to degraded speech intelligibility when the listening task is accompanied by a degree of background noise (79, 143). One of the earliest studies investigating speech intelligibility in background noise was done by Dayal et al. (1966) as part of a series of neuro-otologic studies in pwMS. Discrimination of monosyllables presented with white noise at a signal/noise (S/N) ratio of +10 dB yielded 9 out of the 13 (69%) cases having speech discrimination losses that were disproportionate to their pure tone hearing (143). Reports of normal pure tone thresholds but a loss of speech discrimination are also reported by Noffsinger et al. (1972) in a larger study with 61 pwMS (79). This procedure was administered monoaurally with a S/N ratio of 0 dB at a presentation level of 40 dB of the relevant speech reception threshold (SRT). 37 of the 115 ears in the MS group scored below 60%; the threshold which was denoted to be abnormal (79).

In studies that employed normative data, no significant differences were found in word discrimination scores between treatment groups. Rappaport et al. (1994) presented words with continuous wide-band noise at various S/N ratios and the patient's task was to repeat each word as it occurred. For both controls and patients, word recognition improved with more favorable S/N ratios (i.e. background noise presented at lower intensity than target words). However, there was no main effect of listener type. A later study also yielded no significance when speech discrimination thresholds were determined by a monosyllabic word test assessed at 40 dB in quiet environment and continuous white noise at 0 dB S/N (160).

Speech-in-noise paradigms have been studied in various clinical populations; demonstrating that this type of impaired speech discrimination in the presence of normal pure-tone thresholds is not specific to MS (142). Speech discrimination of spondaic words in quiet and in white noise (0 dB S/N ratio) for patients with cochlear, 8th nerve, brain stem and cortical lesions were also evaluated and compared to MS listeners. Similar to previous findings (86, 160), performance of the MS group and listeners with noise trauma were not significantly poorer than controls. However, the scores obtained from some pwMS were poorer than scores for any of the normal participants: 14% of pwMS had abnormal differences of 40% or more between scores in quiet and white noise; despite normal hearing sensitivity for all audiometric test frequencies and excellent speech discrimination in quiet. Ears with peripheral involvement such as Meniere's disease and 8th nerve tumors yielded higher abnormal performances with percentages of 48% and 62% respectively (142).

In addition to the speech-in-noise paradigm involving words being presented against a continuous wide-band noise background, Rappaport et al. (1994) also utilized a background noise that was also wide band, but interrupted by randomised silent periods between. With this type of background stimulus, pwMS had significantly lower word recognition scores than control participants. In normal adults tested with this paradigm, the performance-intensity curves are roughly sigmoidal in shape and performance was considerably superior to discrimination against the continuous masker – especially at unfavourable S/N ratios. This superiority was attributed to the normal listener's ability to use the gaps in the background to resolve the brief speech fragments that occur between them. No significant interaction between the two treatment groups indicated that pwMS were still able to exploit the use of these gaps- but not to the same degree as controls (86).

The aforementioned speech-in-noise tasks have all employed similar white-noise energetic maskers and words or syllables as target speech and thereby, lack the ecological relevance to inform how pwMS function in daily life. Natural sentences in the presence of multiple competing speech sounds represent more ecologically relevant listening situations and is likely to have a more adverse masking effect as it contains greater perceptual masking (139, 161, 162). How pwMS perceive natural speech in the presence of competing talkers remains to be investigated.

1.9 Conclusion

Despite hearing loss being a rare event, pwMS are likely to display abnormalities in various central auditory tasks, depending on the CNS site where damage has occurred. The most established and prominent auditory deficits described in MS implicate the temporal auditory domain. Disruptions in normal functioning of temporal mechanisms are likely to interfere with daily speech processing as speech is a complex sound consisting of rapid spectro-temporal fluctuations, rich harmonic structures, and dynamic amplitude modulations that the brainstem must faithfully encode (163-165). How this might translate to the ability of an MS individual to communicate and function in a crowded world of obstacles, objects and other individuals remains to be reported.

Preface – Chapter Two

An overview of the literature in **Chapter One** highlights the widely accepted view that standard pure-tone audiogram and speech reception threshold tests do not provide a comprehensive or informative assessment of auditory function in multiple sclerosis (MS). This clinical population is likely to display deficits in auditory measures involving central auditory processing, thereby, we hypothesized that people with MS (pwMS) were likely to display deficits in discriminating speech from noise. This was formally tested using a battery of speech-in-noise (SiN) tasks and a questionnaire designed to reflect auditory processing behaviours in daily life.

The study presented is written up as a manuscript for publication and submitted for review. Formatting has been changed to maintain consistency with other chapters, however some formatting inconsistencies will still be present due to the requirements of the journal.

Chapter Two: Speech Discrimination Impairments as a Marker of Disease Severity in Multiple Sclerosis

Pippa Iva[a], Joanne Fielding[b], Meaghan Clough[b], Owen White[b], Gustavo Noffs[c], Branislava Godic[a], Russell Martin[a], Anneke van der Walt[b], Ramesh Rajan[d]

[a]Department of Physiology, Monash University, Melbourne, Australia

[b]Department of Neurosciences, Central Clinical School, Alfred Hospital, Monash University, Melbourne, Australia

[c]Centre for Neuroscience of Speech, University of Melbourne, Melbourne, Australia

[d]Neuroscience Program Biomedicine Discovery Institute, School of Biomedical Sciences, Monash University, Melbourne, Australia

Corresponding author details:

Pippa Iva

pippa.iva@monash.edu

26 Innovation Walk, Monash University, Clayton, Victoria, Australia 3168

+61 420 653 560

Abbreviations

MS = multiple sclerosis; SNR = signal-to-noise ratio; EDSS = Expanded Disability Status Scale; dB = decibel; pwMS = people with multiple sclerosis; AEP = auditory evoked potential; SiN = speech-in-noise; HREC = Human Research Ethics Committee; STROBE = Strengthening the Reporting of Observational Studies in Epidemiology; Hz = hertz; dB HL = decibel hearing level; ISO = International Organization for Standardization; BKB = Bamford-Kowl-Bench; SWN = speech-weighted noise; BN = babble noise; AADQ = Auditory Attention and Discomfort Questionnaire; RR = relapsing-remitting; SP = secondary progressive; VIF = variance inflation factor; AIC = Akaike information criterion; LRT = likelihood ratio test; CAPD = central auditory processing disorder

Abstract

Background:

Multiple sclerosis (MS) pathology is likely to disrupt central auditory pathways, thereby affecting an individual's ability to discriminate speech from noise. Despite the importance of speech discrimination in daily communication, it's characterization in the context of MS remains limited. This cross-sectional study evaluated speech discrimination in MS under "real world" conditions where sentences were presented in ecologically valid multi-talker speech or broadband noise at several signal-to-noise ratios (SNRs).

Methods:

Pre-recorded Bamford-Kowal-Bench sentences were presented at five signal-to-noise ratios (SNR) in one of two background noises: speech-weighted noise and eight-talker babble. All auditory stimuli were presented via headphones to control (n=38) and MS listeners with mild (n=20), moderate (n=16) and advanced (n=10) disability. Disability was quantified by the Kurtzke Expanded Disability Status Scale (EDSS) and scored by a neurologist. All participants passed a routine audiometric examination.

Results:

Despite normal hearing, MS psychometric discrimination curves which model the relationship between signal-to-noise ratio (SNR) and sentence discrimination accuracy in speech-weighted noise and babble did not change in slope (sentences/dB) but shifted to higher SNRs (dB) compared to controls. The magnitude of the shift in the curve systematically increased with greater disability. Furthermore, mixed-effects models identified EDSS score (odds ratio = 0.81; $p < 0.001$) as the most significant predictor of speech discrimination in noise. Neither age, sex, disease phenotype or disease duration were significantly associated with speech discrimination in noise. Only MS listeners with advanced disability self-reported audio-attentional difficulty in a questionnaire designed to reflect auditory processing behaviours in daily life.

Conclusion:

Speech discrimination performance worsened systematically with greater disability, independent of age, sex, education, disease duration or disease phenotype. These results identify novel auditory processing deficits in MS and highlight that speech discrimination tasks may provide a viable non-invasive and sensitive means for disease monitoring in MS.

2.1 Introduction

Multiple sclerosis (MS) has a heterogeneous clinical course and symptomology which includes disruption of motor, cognitive and sensory systems(7). Despite the importance of hearing to communication, characterization of auditory deficits in people with MS (pwMS) remains inconsistent and elusive(10). Auditory processing networks are highly integrated and widespread(166) and MS-related neurogenic injury at any anatomical level(s) will impact on the person's ability to navigate the world, build relationships, and socialize, all directly impacting quality of life(167).

Although cochlear hearing loss is uncommon in pwMS(81), reports of abnormal wave amplitude and latency in auditory evoked potentials (AEP) from brainstem, subcortical and cortical regions are common(168). Tasks involving later stages of auditory processing, especially psychoacoustic tasks that require binaural hearing and precise neural timing(124, 126), can also be impaired. Binaural hearing requires the listener to integrate complementary sound inputs to both ears, and involves detection within millisecond precision; a function particularly susceptible to effects of demyelination on neural timing and conduction velocity(10). Binaural hearing is vital to many everyday dynamic functions like identifying the location and direction of a sound source (an approaching vehicle), segregating different streams of auditory information (the ringing of a phone from background music), suppressing interference from echoes and reverberations, providing situational awareness (obstacles like workmen drilling in the path), and disambiguating speech in noisy environments(169). We therefore postulate MS individuals will have speech-in-noise (SiN) processing deficits.

How speech processing deficits, especially in real-world conditions, contribute to communication difficulties in MS is poorly studied; communication breakdown in MS is generally reported only in the context of speech production, i.e dysarthria(170). This may be due to observations that MS individuals do not have problems repeating speech presented in silence(142, 143). However, studies attempting to understand real-life communication should consider the fact that our world is often noisy. Descriptions of speech processing in noise(130, 142, 160) are yet to be described in an ecologically relevant context, or are limited in characterizing disease severity in MS listeners. Hence, our objective in this cross-sectional study was to evaluate speech processing in MS under a "real world" perspective of open-set whole sentences in ecologically valid multi-talker speech ("babble") or broadband noise. To examine the impact of disease severity on these functions, pwMS were segregated according to their Expanded Disability Status Scale (EDSS) score(171), a clinical scale widely used for assessing physical disability in MS. Objective test data were also compared to responses in a questionnaire(141) for self-reports of difficulties in different daily-life scenarios.

2.2 Methods

All procedures were approved by the Monash University Human Research Ethics Committee (8170) and Melbourne Health HREC (2015.069). The study conformed to guidelines of the National Health and Medical Research Council of Australia and the Helsinki Declaration protocols for experiments involving human participants.

2.2.1 Participants

Forty-six people with confirmed MS by revised McDonald criteria(172) were recruited through the Royal Melbourne Hospital Australia. Thirty-eight controls without MS were recruited from the local community. All participants provided informed written consent. The main exclusion criterion for all participants was hearing loss (section 2.3 Audiometry for definition of hearing loss), confirmation of no other neurological disorder (with the exception of MS for the patient group), and no recent (within 30 days) relapses and/or steroids administration in the case of pwMS. All participants reported English as their native language.

PwMS were grouped according to EDSS score (171) as rated by a neurostatus certified neurologist at study entry. PwMS with EDSS scores ≤ 1.5 were classified as 'mild'; between 2 – 4.5 as 'moderate' and between 5 – 7 as 'advanced' disability.

2.2.2 Study Overview

This study was designed, implemented, and reported in accordance with the Guidelines for the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)(173). All participants completed an assessment battery of audiometry, speech discrimination tasks, and an auditory questionnaire, in a quiet room over a single session lasting 35-45 minutes.

2.2.3 Pure-tone audiometry

Hearing status was determined using a Beltone Model 110 Clinical Audiometer and calibrated TDH headphones to test sensitivity one ear at a time, at standard audiometric frequencies of 250 Hertz (Hz), 500Hz, 750Hz, 1000Hz, 1500Hz, 2000Hz, 4000Hz, 6000Hz and 8000Hz, using a modified Hughson-Westlake procedure(174). Hearing thresholds, recorded as decibels Hearing Level (dB HL) relative to normal sensitivity (ISO 8253-1, 1989), were defined as the lowest level at which the tone was perceived 50% of the time. Pure tone averages (PTAs) of hearing threshold levels at 500, 1000, 2000 and 4000 Hz were obtained for all participants to describe hearing status, and only participants with a bilateral four tone average < 25 dB HL were used in this study. Participants with hearing loss (≥ 25 dB) were excluded to remove peripheral hearing loss as a confounding factor on speech discrimination ability.

2.2.4 Speech-in-noise discrimination tasks

The general procedures and stimuli for the SiN task have been detailed previously(141, 175). In brief, speech stimuli, derived from a standard battery of clinically used sentences called the Bamford-Kowal-Bench (BKB) sentence

lists(176), were each four to six words long with three keywords (**Supplementary Figure A.1**). Sentences were presented in speech-weighted noise (SWN) or babble noise (BN). SWN was shaped to the long-term average spectrum of the target sentences(141). BN consisted of eight simultaneous voices generated by doubling over and temporally offsetting a recording of four people reading nonsense text.

Speech and masker stimuli were presented binaurally through Sennheiser HD535 headphones. Sentences were presented at a constant level of 70 dBA, whilst the masker level was varied to generate signal-to-noise ratio (SNRs) of 3, 1, -1, -3, and -5 dB in BN; and 1,-1,-3,-5,and -7 dB in SWN. Prior to each noise condition, participants completed ten practice trials (ten unique target sentences) at an 'easy' SNR of +5 dB for acclimatization to stimuli. Subsequent SNR blocks were presented in random order. At each SNR, ten unique sentences were presented one at a time and the listener asked to repeat each sentence or indicate inability to do so. A correct response was scored when all three keywords were correctly repeated in correct order. No time limit was placed on response and feedback was not provided. The experimenter recorded the responses and presented the next sentence after 1.5 second delay.

2.2.5 Auditory attention and discomfort questionnaire

The auditory attention and discomfort questionnaire (AADQ) was developed by Dunlop, Enticott and Rajan (2016) and based on validated inventories for specific adult clinical populations with abnormal auditory processing(141). The 33-item AADQ consisted of statements about daily life events involving hearing and had three subscales; the Audio-Attentional Difficulty subscale measured difficulties attending to speech in noisy environments; the Auditory Discomfort (Non-Verbal) subscale measured discomfort to non-verbal environmental sounds; and the Auditory Discomfort (Verbal) subscale measured discomfort to verbal sounds.

2.2.6 Generalised linear mixed model

To identify factors that significantly influenced speech discrimination accuracy on any given trial (0=incorrect; 1=correct), two binomial generalised linear mixed effects model with logit link functions were generated using MATLAB Statistic Toolbox Release 2019b. One model was based on all participants, whilst the other focused on pwMS only. To build the models, considered variables included: disability groups, SNR, masker type (SWN vs. BN), trial order, age (years), sex (male vs. female), education (years), average pure-tone thresholds (dB HL), EDSS score, disease duration (years), disease phenotype (relapsing-remitting (RR) vs. secondary-progressive (SP)) and theoretically relevant interactions. **Supplementary Table A.1** specifies how the categorical/ordinal variables were coded, and the mean ± SD and range for continuous variables. Potential fixed-effects were explored with a participant-specific random intercept representing between-participant heterogeneity. All variables had variance inflation factors (VIF) < 3, below the recommended cut off VIF of 5, indicating no problematic levels of multicollinearity among predictors.

Models were validated using the 'hold-out method', with a 70:30 split into training and validation data sets, and confusion matrices were generated to determine sensitivity (true positive rate) and specificity (true negative rate).

2.3 Results

2.3.1 Participant groups

Basic demographics and disease details of the participant groups are reported in **Table 1**. Twelve controls (24%) and eleven MS (26%) participants were excluded for bilateral hearing loss (thresholds ≥ 25dB HL) at some/all the frequencies between 250 – 4000 Hz. The remaining participants (**Table 1**) had bilaterally normal hearing between 250 – 4000 Hz except for 5% of participants from each group with small hearing losses (of 5–10dB) at higher frequencies of 6000 and 8000 Hz in one ear only. Controls and MS groups had comparable hearing sensitivity (see **Supplementary Figure A.2**).

Table 1. Participant characteristics

	Control	Mild MS	Moderate MS	Advanced MS
Number of participants	38	20	16	10
Sex F(M)[ˆ]	35(3)	17(3)	13(3)	9(1)
Phenotype RR(SP)[ˆˆ]	-	20(0)	13(3)	2(8)
Age (yrs)				
Mean (SD)	45.66(10.43)	44.3(9.52)	44.83(11.69)	49(6.56)
Range	28 - 60	24 - 63	28 - 64	36 - 58
Disease duration (yrs)				
Mean (SD)	-	10.7(5.77)	13(7.24)	18.5(7.07)
Range		1 - 22	1 - 32	10 - 31
EDSS*				
Median	-	0	2.5	6
Range	-	0-1.5	2-4.5	5-7
Disease modifying therapy (%)	-	90	81	80

*EDSS = Expanded Disability Status Scale Score determined by a neurologist within 6 months of audiological testing.
[ˆˆ]RR = Relapsing-remitting; SP = Secondary progressive;
[ˆ] F = female; M = male.

2.3.2 Speech discrimination in noise

In both noise conditions, sentence recall decreased as SNR decreased. SiN discrimination appeared to be easier in SWN than BN as a floor effect occurred in BN at an SNR of -5, at which point sentence recall was poor for all listener groups, however, at the same SNR of -5 in SWN, no such floor effect was observed. A direct comparison between the noise conditions is described in a mixed effects model described in **Section 3.4**.

2.3.3 Identification of sentences in speech-weighted noise

Mean ± SEM sentences in SWN correctly recalled by controls and pwMS at various SNRs is presented in **Figure 1A**. A 4 x 5 [(control, minimal, moderate and advanced MS)] x (SNR = 1, -1, -3, -5, and -7)] two-way mixed ANOVA confirmed a significant interaction between listener group and SNR on sentence recall [$F(12, 308)=2.45, p=.005$]. There was also a significant main effect for listener group [$F(3,77)=16.66, p<0.0001$] and SNR [$F(4,308)=372.1, p<.0001$]. A Tukey's post hoc analysis confirmed that significantly fewer sentences were recalled by moderate ($p=.0004$) and advanced ($p<.0001$), but not mildly impaired pwMS ($p=.46$) compared to controls.

2.3.4 Identification of sentences in multi-talker babble

BN degraded speech intelligibility for all MS listener groups more than controls except at an SNR of -5, at which a floor effect was observed (refer to **Figure 1B**). A 4 x 5 two-way mixed ANOVA confirmed a significant interaction effect between listener group and SNR [$F(12, 320)=3.445, p<.0001$]. Main listener group effects were also significant [$F(3, 80)=16.86, p<.0001$]; and as expected, the SNR also had a significant effect on sentence recall in BN [$F(4,320)=595.6, p<.0001$]. A Tukey's post hoc analysis confirmed that significantly fewer sentences were discriminated by all MS listener groups ($p<0.05$) compared to controls.

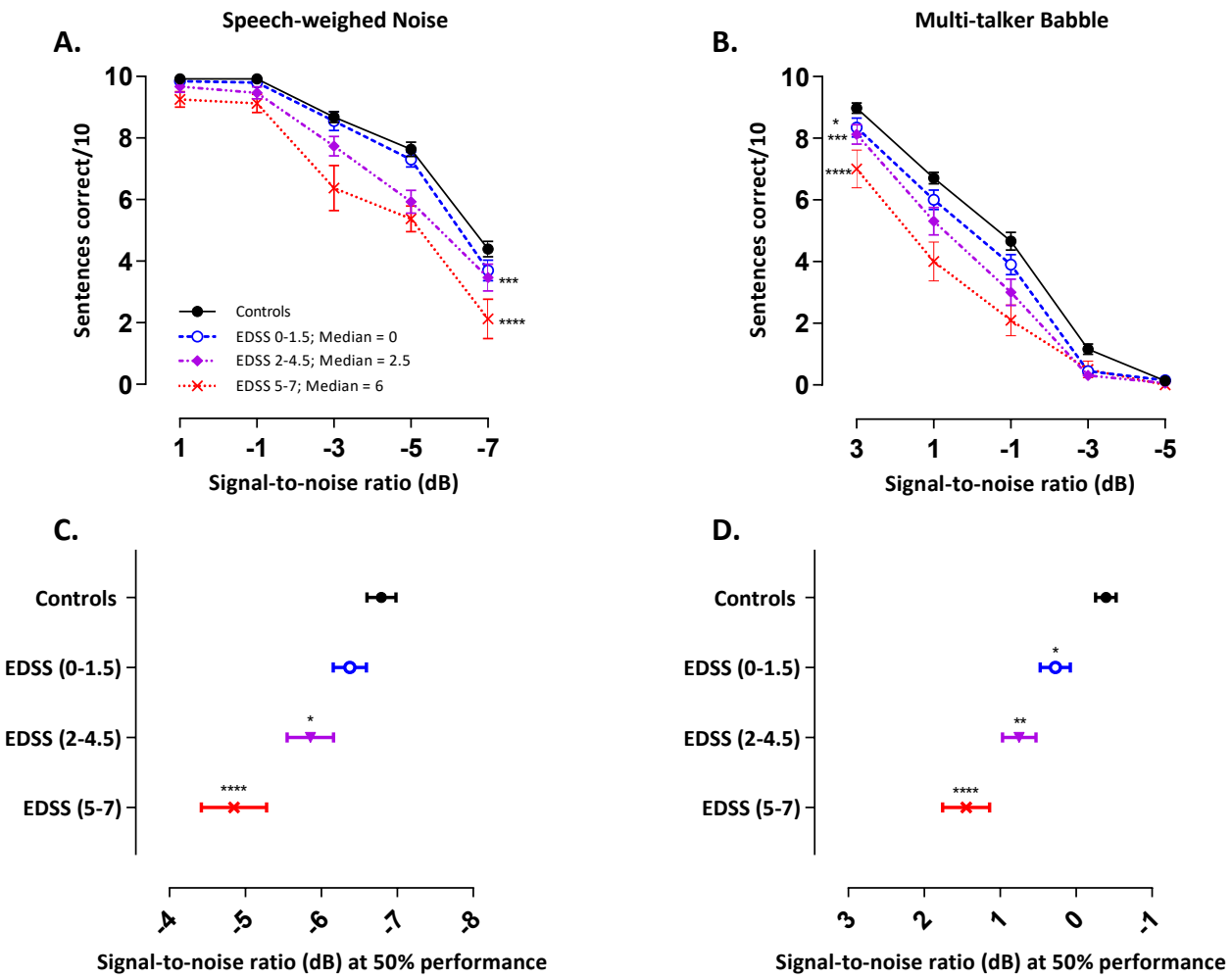


Figure 1. Sentence recall was systematically worse in pwMS with greater disease severity as measured by the Expanded Disability Status Scale (EDSS) score. Sentence recall by controls (filled circle, n = 38); and pwMS with mild (EDSS 0–1.5, open circle, n=20); moderate (EDSS 2–4.5, diamond, n=16); and advanced disability (EDSS 5-7, cross, n=10) in speech-weighted noise (SWN; left-hand column) and multi-talker babble (BN; right-hand column). Mean ± SEM sentences correctly discriminated (/10 at each signal-to-noise ratio) in SWN **(A)** and BN **(B)**. Mean SNRs ± SEM (dB) at 50% discrimination in SWN **(C)** and BN **(D)**. * $p < 0.05$; *** $p < 0.001$; ****$p< 0.0001$ compared to controls. (**A&B**: 2-way ANOVAs; **C, D**: 1-way ANOVA).

2.3.5 Estimating psychometric functions

To quantify MS effects on SiN discrimination, Boltzmann sigmoidal functions were fitted to each participant's discrimination curves, using GraphPad PRISM 6. From each psychometric curve the slope and midpoint data were extracted (see **Supplementary Tables A.2 & A.3** for details). A one-way ANOVA revealed no significant difference in slopes (sentences/dB) between the listening groups in SWN [$F(3,77)=1.70, p=0.18$] and BN [$F(3,80)=0.3, p=0.83$]. In contrast, the midpoints of the curves were significantly different amongst listener groups in SWN [$F(3,77)=7.48, p=0.0002$] and BN [$F(3,80) = 14.84, p<0.0001$]. The midpoints represent the SNR±SEM (dB) at 50% discrimination and are visually graphed for the SWN (**Figure 1C**) and BN task (**Figure 1D**), note: higher SNRs indicated poorer discrimination performance.

A Tukey's honestly significant difference (HSD) post hoc test confirmed that in SWN, the SNR at 50% discrimination for moderate and advanced pwMS was significantly higher than controls ($p<0.05$). However, no statistical difference was found between controls and minimally impaired pwMS ($p=.55$). In BN, the SNR at 50% discrimination for all MS groups was significantly higher than controls ($p<0.05$). Minimal, moderate, and advanced pwMS had 0.7±0.35 dB, 1.14±0.26 dB and 1.84±0.31 dB greater SNRs than controls, respectively.

2.3.6 Modelling the factors that impact on SiN discrimination

To explain the impact of MS on SiN discrimination, we adopted a holistic approach to build a model that incorporates all variables needed for explanatory power. Model-building started with a 'constrained model' with fixed effects being disability group, SNR, masker type, and the interaction between SNR and masker. Additional theoretically important variables such as trial order number, age, sex (male vs. female), education (years) and average pure-tone thresholds (dB HL) were then incorporated and the model evaluated (see **Supplementary Table A.1** for details on how variables were coded). Thirteen theoretical regression models (**Table 2**) were generated using MATLAB Statistic Toolbox Release 2019b and compared to the constrained model to determine the difference in Akaike's Information Criterion (AIC:ΔAIC). Models were based on 8250 trials as nearly all participants (N=84) completed 100 trials each (10 trials x 5 SNRs x 2 masker types). The model with the lowest AIC was used to select the final model as it explains the greatest amount of variation using the fewest possible independent variables. Tests of fixed effects were also confirmed with likelihood ratio (LRT) tests to compare the constrained model with nested models.

The addition of demographic variables did not significantly improve the constrained model. The model with the lowest AIC value included the addition of 'trial order' as a fixed effect (**model 6** in bold in **Table 2**). Therefore, the constrained model was rejected in favour of the final model 6.

Table 2. Comparisons of fixed effects combinations in a generalised linear mixed effects model (with a logit link function) used to predict correct sentence recall on each trial

Constrained model (nested model):	***AIC***	***ΔAIC***	$\mathbf{X^2}$	***Δdf***	***p***
(1) Disability group + SNR + Masker + SNR*Masker + (1|participant)	6615.7	0			
Additional covariates:					
(2) Disability group*SNR	6618.3	2.6	3.37	3	0.34
(3) Disability group*Masker	6619.9	4.2	1.77	3	0.62
(4) Disability group*SNR + Disability group*Masker	6621.1	5.4	6.59	14	0.36
(5) Disability*SNR*Masker	6624.3	8.6	9.34	9	0.41
(6) Trial order	**6610.3**	**-5.4**	**7.38**	**1**	**0.007****
(7) Trial order + Trial order*Disability group	6615.3	-0.4	8.42	4	0.08
(8) Trial order + Age + Sex + Education + Pure tone average	6617.1	0.4	7.41	5	0.21
(9) Pure tone average	6617.5	1.9	0.08	1	0.78
(10) Pure tone average*Disability group	6617.6	0.9	7.10	4	0.13
(11) Age	6617.7	2	0.01	1	0.92
(12) Age*Disability group	6622.4	6.7	1.31	4	0.86
(13) Sex	6617.8	2	0.05	1	0.82

Based on 8250 observations (84 participants)
The estimate of the variability of the random effects ($\boldsymbol{\sigma_b^2}$) for all models = 0.32.
The last three columns show the chi squared statistic (X^2), difference in the degrees of freedom and p value from the likelihood ratio (LRT) tests
SNR = signal-to-noise ratio
AIC = Akaike's Information Criterion
df = degrees of freedom

Parameter estimates of fixed effects in model 6 are listed in **Table 3**, along with the t statistic, degrees of freedom and p-values for each fixed effect to test the null hypothesis that the coefficient equals zero. The odds ratio (OR) and confidence intervals (CI,95%) are also included to quantify the magnitude of the association between the fixed effect and the outcome.

Changes in all fixed effects are significantly associated with changes in sentence recall accuracy ($p<0.05$). To quantify this association, the ORs were interpreted. The OR of correctly discriminating a sentence was 0.72(0.56 – 0.92) for mildly impaired pwMS, i.e 28% lower odds (1- $e^{-0.33}$) compared to controls when all other factors were constant. A decrease in OR corresponded to the severity of disability as moderately and advanced pwMS had 51% and 73% lower odds compared to controls, respectively.

There was also a positive association of trial order on speech discrimination as the odds of correct speech discrimination was 1.03 (95% CI, 1.01-1.06) times greater compared to the previous trial. Finally, a one-unit increase in SNR increased the odds of correct sentence recall by 84% in SWN but only by 11% in BN.

Table 3. Parameter estimates of fixed effects of the final generalised linear mixed effects model (with a logit link function) used to predict correct sentence recall on each trial

						95% C.I for e^{β}	
Name	**Estimate (β)**	**SE β**	**tStat**	***P***	**OR (e^{β})**	**Upper**	**Upper**
Intercept + $N(0, \sigma_b^2)$	3.88	0.15	26.15	**<0.0001**	48.26	36.09	64.54
Trial	0.03	0.01	2.72	**<0.01**	1.03	1.01	1.05
SNR	0.59	0.02	27.08	**<0.0001**	1.81	1.73	1.89
Masker type	-3.86	0.12	-30.94	**<0.0001**	0.02	0.02	0.03
Disability group							
Control	Reference group		-	-	1	-	-
Mild (EDSS 0-1.5)	-0.33	0.12	-2.66	**<0.01**	0.72	0.56	0.92
Moderate (EDSS 2-4.5)	-0.71	0.13	-5.27	**<0.0001**	0.49	0.38	0.64
Advanced (EDSS 5-7)	-1.30	0.16	-7.92	**<0.0001**	0.27	0.20	0.38
SNR x Masker type	0.11	0.03	3.62	**<0.0001**	1.11	1.05	1.18

Based on 8250 observations (N = 84 participants)
The estimate of the variability of the random effects (σ_b^2) for all models = 0.32
Model was validated using the 'hold-out method', with a 70:30 split into training and validation data sets
Note: Sensitivity = 82.7%. Specificity = 80.5%. Overall accuracy = 81.25%
SNR = signal-to-noise ratio
SE = standard error
OR = odds ratio
C.I = confidence interval

2.3.7 Disease factors that impact SiN discrimination

A second model was built to investigate the impact of various disease factors on SiN discrimination, this time, with a specific focus on characteristics and clinical measures in pwMS only. In an exploratory approach, all theoretically important variables were included. The model was based on 4450 trials as nearly all participants (N=46; with the exception of one participant who did only 50 trials) completed 100 trials each (10 trials x 5 SNRs x 2 masker types). A total of 11 variables were identified for inclusion in the generalised linear mixed effects model, and were classified into four groups: experimental, demographic, disease characteristics and 'other'. Experimental fixed-effects were: trial order, SNR, masker type, and the interaction between SNR and masker; demographic variables were: age (years), sex (female: male) and education (years); disease characteristic variables were: duration (years), EDSS score and disease type (RR:SP); and the 'other' variable was: the average pure-tone threshold (dB HL). Refer to **Supplementary Table A.1** for details on how variables were coded.

Figure 3 displays the ORs for the demographic, disease characteristics and pure-tone average variables that were included in the model. Although the fixed effects of trial order, SNR, masker, interaction between SNR and masker, and the Intercept were significant predictors of the model, they are not displayed in **Figure 3** as it was not informative to repetitively display experimental variables that have previously been established as significant contributors to speech discrimination accuracy (refer to **Table 3**). EDSS score (OR 0.81; $p < 0.001$) was the only predictor identified as the most significant predictor of speech discrimination in noise. There were no associations between any of the other patient characteristics and SiN discrimination.

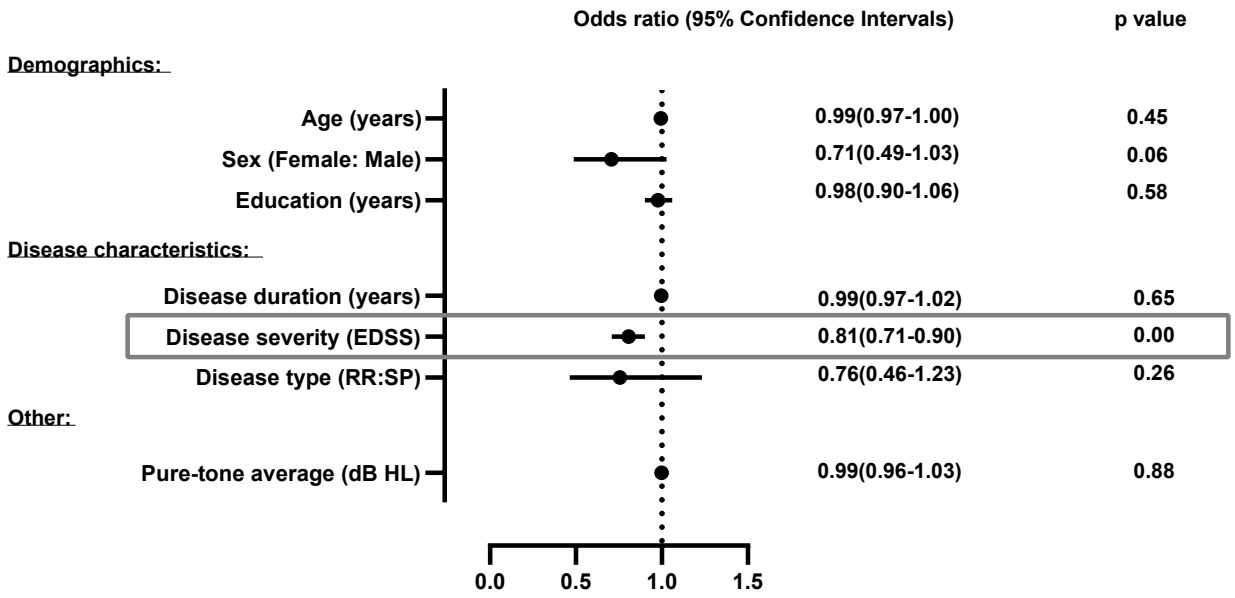


Figure 3. Disease severity, as measured by the Expanded Disability Status Scale (EDSS) score, was the only significant predictor for correct speech discrimination in noise at any given trial. A generalised linear mixed-effects model (with a logit link function) was used to determine significant predictors of speech discrimination in people with Multiple Sclerosis; the grey box highlights the significant predictor ($p < 0.05$). An odds ratio of 1 = no effect; < 1 is associated with lower odds of correctly discriminating a sentence from noise. EDSS: Expanded Disability Status Scale Score; RR: Relapsing Remitting; SP: Secondary Progressive; dB HL: decibels hearing level. Black dots indicate odds ratio and lines indicate the 95% confidence interval. Refer to Supplementary Table A.1 for details on how variables were coded.

2.3.8 Noise and daily life events

Data for the three AADQ domains probing subjectively perceived difficulty in different facets of daily life events are presented in **Figure 4**. ANOVA revealed significant differences between groups on Audio-Attentional difficulty [$F(3, 77) = 7.05$, $p = .0003$; $\eta^2 = .22$], but no significant differences on the Auditory Discomfort scales for both non-verbal [$F(3, 77) = 1.30$, $p = .28$, $\eta^2 = .05$] and verbal stimuli [$F(3, 77) = 2.09$, $p = 0.11$; $\eta^2 = .08$]. Tukey's post-hoc analysis confirmed that only the advanced MS group reported significantly greater difficulty in attentionally demanding environments than controls ($p < .001$), mild ($p < .01$) and moderately impaired ($p < .05$) pwMS.

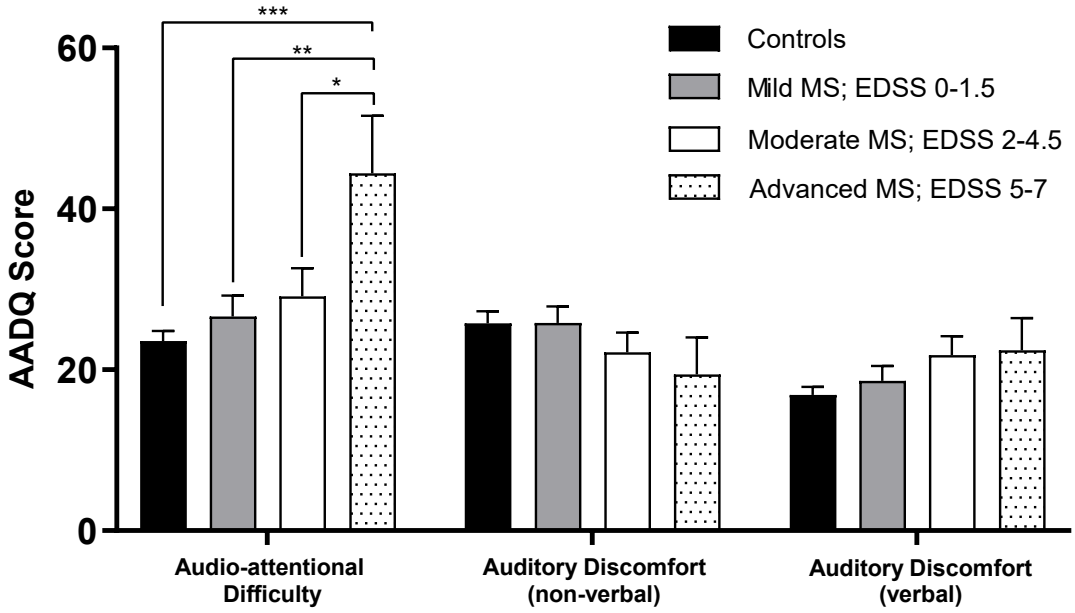


Figure 4. PwMS with advanced disability reported significant audio-attentional difficulty during daily life events. Mean total score (±SEM) for controls (black; n=38), mild (grey; n=20), moderate (white; n=15) and advanced MS (patterned; n=7) on the three components of the questionnaire: audio-attentional difficulty, auditory discomfort (non-verbal) and auditory discomfort (verbal). Audio-Attentional Difficulty had a possible range of 14-98, Auditory Discomfort (Non-Verbal) had a possible range of 8-56, and Auditory Discomfort (Verbal) had a possible range of 5-35. (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; One way ANOVA; Tukey's Honestly Significant Difference (HSD) Test).

2.4 Discussion

We have uniquely investigated sentence discrimination under 'simulated real-world' conditions to report deficits in MS listeners in discriminating open-set natural whole sentences in noise, including ecologically valid noise. This worsened systematically with disease severity but the absent slope changes coupled with the shift to higher SNRs for 50% performance in MS psychometric curves show that MS groups required a more favourable SNR for equal performance but otherwise conducted SiN processing in the same way as control listeners.

Successful extraction of speech from noise requires disentangling a complex auditory scene to develop neural representations that maintain the integrity of distinct sound sources(177). Given audiometric normal hearing, MS-related SiN difficulties must reflect centralised auditory processing disorders (CAPD) (10) in higher-order mechanisms that preserve, analyse, organise and interpret information. Temporal processing, an important component of central auditory processing, is impaired in MS(86); posited to be related to the delays of signal transmission within the auditory pathways affected by demyelinating lesions that impact on neural synchrony(11, 86). Temporal acuity is critical for speech perception processes like detection of rise/fall time, voice onset, and the transient onset of syllables(178), and it also facilitates 'glimpsing' in which an individual takes advantage of momentary 'dips' in noise energy where the target signal is more audible(179). This ability appears impaired in MS listeners who have been previously reported to perform worse than controls in a words-in-noise paradigm when a wideband background noise had randomised silent periods, but not for continuous noise(86).

Sound input parsed by spectral/temporal cues is modified at cortex to sharpen stream segregation by attentional systems that filter irrelevant inputs so the listener can focus on a single target stream(177). This becomes more difficult when SNRs are smaller and stimuli are similar. Multi-talker babble will elicit confusion because of its similarity to speech and its saliency which will involuntarily capture attention. This perceptual interference is known as informational masking and similarity between talkers is a particularly strong feature of such masking(180). In contrast, SWN is an energetic masker that diminishes target audibility only through masking and blending of acoustic signals at the periphery(180). The difference in difficulty of the two tasks is apparent in our modelling: when all other factors were constant, a one-unit improvement in SNR increased the odds of correct sentence recall by 84% in SWN but only 11% in babble. However, there was no differential degradation between the disability groups.

Electrophysiology shows MS-related impairments in cortical processes that can affect discrimination performance. The cognitive P300 potential is elicited in central processes such as attention, auditory discrimination, memory and decision making(181). It is typically measured with an "Oddball Paradigm" requiring a response to deviants within a regular train of repetitive stimuli. Parallels can be drawn to our SiN tasks: background noise forms the repetitive stimulus and target sentences are the deviants requiring detection. In MS, P300 waveform latencies are significantly increased (>2 S.D)(108, 109), indicating poorer cognitive performance (181). In fact, cognitive impairment is now considered a primary deficit affecting 40-70% of pwMS, manifesting at all disease stages, even onset, and in all subtypes(182). MS affects many cognitive domains with most effects on information processing

speed, attention and memory, followed by verbal fluency and executive deficits(7); cognitive decline worsens with advancing disease. Such cognitive disturbances in MS could contribute to impaired SiN discrimination, however, were not formally tested.

Although SiN discrimination was impaired in all MS groups, only pwMS with advanced disability reported significant audio-attentional difficulty in daily life events. The absence of self-reported auditory difficulty in less severe MS groups could reflect redundant auditory processing(10), which may be intrinsic (multiple parallel auditory CNS representations(166)) or extrinsic (syntactic and semantic cues, or multimodal information through (say) lipreading)(183). Early pwMS may successfully use compensatory mechanisms to reduce or mask functional deficits(67). Disease progression may degrade compensatory capacity by causing irreversible neurological disability and whole brain volume atrophy(184), removing any auditory pathway redundancy. Our subjective measures and psychoacoustic testing serve as complementary tasks to elucidate the difference between a subtle impairment that evades detection and one that greatly impacts on daily life.

Our generalised linear mixed effects model enabled consideration of demographic variables and individual differences (random intercept effect) inherent in human participant trials but even then, disease severity remained a significant factor in predicting speech discrimination accuracy. Thus, our SiN tasks have robust construct validity and merit consideration for evaluating disease burden, with the advantages of speed (approximately 10 minutes per background noise) and being non-invasive, cost effective, easy to administer, and requiring only portable equipment, allowing for home testing. Furthermore, psychoacoustic methodology makes it easy and cost effective to study many more systematic variations in SNR, sentence difficulty and saliency of background maskers for further refinement. Finally, we acknowledge that our study is cross-sectional and limited to participants with normal hearing. Longitudinal data will provide further confidence that our SiN tasks could be a valid biomarker for disease progression and future studies should investigate SiN performance in pwMS with hearing loss. An investigation into the correlations between SiN performance and CNS lesion location would also provide valuable insight into the pathological underpinnings of SiN deficits in MS.

Preface – Chapter Three

Chapter Two established that people with multiple sclerosis (pwMS) have impaired speech discrimination in noise which is worse in those with greater physical disability. This current chapter will now focus on the clinical utility of speech-in-noise (SiN) tasks in patients with early and mild disease that typically present with subtle disease changes that are overlooked in the clinic. Quantifying functional deficits in early/mild disease is necessary to maximize the opportunities for clinicians to preserve neurological reserve in patients with appropriate therapeutic management.

* It must be noted that, whilst MS groups in this study are grouped according to disease duration, the purpose of this study is to determine if SiN measures can quantify neurological deficits in participants with early, as well as mild disease - not to investigate effects of disease duration as the multi-regression analysis in **Chapter Two** confirmed that duration is not a factor that influences SiN performance.

The study presented is written up as a manuscript for publication and submitted for review. Formatting has been changed to maintain consistency with other chapters, however some formatting inconsistencies will still be present due to the requirements of the journal.

Chapter Three: Speech Discrimination Tasks: A Sensitive Sensory and Cognitive Measure in Early and Mild Multiple Sclerosis

Pippa Iva[1], Joanne Fielding[2], Meaghan Clough[2], Owen White[2], Branislava Godic[1], Russell Martin[1], Ramesh Rajan[3]
[1]Department of Physiology, Monash University
[2]Department of Neurosciences, Central Clinical School, Alfred Hospital, Monash University
[3]Neuroscience Program Biomedicine Discovery Institute, School of Biomedical Sciences, Monash University

ABSTRACT

Background: There is a need for reliable and objective measures of early and mild symptomology in multiple sclerosis (MS), as deficits can be subtle and difficult to quantify objectively in patients without overt physical deficits. We hypothesized that a speech-in-noise (SiN) task would be sensitive to demyelinating effects on precise neural timing and diffuse higher-level networks required for speech intelligibility, and therefore be a useful tool for monitoring sensory and cognitive changes in early MS.

Objective: The objective of this study was to develop a SiN task for clinical use that sensitively monitors disease activity in early (<5 years) and late (>10 years) stages of MS participants with mild severity (Expanded Disability Status Scale (EDSS) score <3).

Methods: Pre-recorded Bamford-Kowal-Bench sentences and isolated keywords were presented at five signal-to-noise ratios (SNR) in one of two background noises: speech-weighted noise and eight-talker babble. All speech and noise were presented via headphones to controls (n = 38), early MS (n = 23) and late MS (n =12) who were required to verbally repeat the target speech. MS participants also completed extensive neuropsychological testing which included: Paced Auditory Serial Addition Test, Digit Span Test and California Verbal Learning Test.

Results: Despite normal hearing thresholds, participants with early and late mild MS displayed speech discrimination deficits when sentences and words were presented in babble – but not speech-weighted noise. Significant correlations between SiN performance and standardized neuropsychological assessments indicated that MS participants with lower functional scores also had poorer speech discrimination. Furthermore, a quick five-minute task with words and keywords presented in multi-talker babble at an SNR of -3dB was 82% accurate in discriminating mildly impaired MS individuals (median EDSS=0) from healthy controls.

Conclusions: Quantifying functional deficits in mild MS will help clinicians to maximize the opportunities to preserve neurological reserve in patients with appropriate therapeutic management, particularly in the earliest stages. Given that physical assessments are not informative in this fully ambulatory cohort, a quick five-minute task with words and keywords presented in multi-talker babble at a single SNR could serve as a complementary test for clinical use due to its ease of use and speed.

3.1 Introduction

Multiple sclerosis (MS), a debilitating disease of the central nervous system (CNS), is the most common cause of neurological disability in young adults (185). People with MS (pwMS) display a range of motor, sensory and cognitive symptoms that can sometimes cause serious disability, although occasionally can be mild (7, 71). Currently, the gold standard clinical measure of MS disability is the Expanded Disability Status Scale (EDSS) (171). While the EDSS provides a sound measure of motor dysfunction, particularly later in the disease when symptoms are more pronounced, it has been less reliable at detecting early symptomology (EDSS scores of < 3 - early disease, low disability), which are often subtle and difficult to quantify objectively (186, 187). In particular, the EDSS does not characterise cognitive impairment, which is a particularly debilitating component of the disease, affecting 40-70% of pwMS (188) manifesting at all disease stages, including onset (64, 70) and often predating physical symptoms (189). Consequently, there is a need for reliable and objective measures of early MS that are sensitive to the range of symptoms that may occur, particularly for clinical management as evidence suggests that insidious progression during the early phase of MS can meaningfully inform prognostication (190, 191). Furthermore, sensitive measures of disease surveillance also provide a means to evaluate treatment effects of potential and current therapeutics. To monitor early disease activity, we propose an innovative approach using speech-in-noise (SiN) assessments; an auditory process we previously demonstrated to be impaired in pwMS despite normal peripheral hearing (192). SiN is a complex process that integrates sensory information and cognitive processing, that maybe be measured with high sensitivity, allowing a comprehensive measure of sensory and cognitive function in early MS.

Speech is a complex sound consisting of rapidly changing elements that require precise temporal detection within milliseconds for identifying consonants or voice onset time, especially in the presence of background noise (164). Early stages of SiN processing take place subcortically within the auditory brainstem, where binaural (two-ears) sensory integration forms a necessary element of segregating an ambiguous sound mixture into coherent auditory objects (163). Early brainstem involvement is common in MS and accounts for 20% of symptoms supportive of a diagnosis of Clinically Isolated Syndrome (CIS) (193), the earliest stage of disease in 85% of patients who subsequently develop MS (194). In MS, demyelination in the brainstem causes slowed conduction velocity and neural dyssynchrony, resulting in less precision to detect acoustic timing cues within the millisecond range (86, 106). Background noise simultaneously presented with speech further degrades neural synchrony by disrupting the representation of temporal characteristics of the stimulus (164, 195), potentially making SiN assessments a potent measure for detecting MS changes. It's possible that such MS-induced deficits in processing timing cues create significantly more degradation of the target speech compared to healthy controls, which places a greater cognitive load on the MS listener to use top-down processes to increase speech intelligibility (196-199). As proposed in the ease-of-language-understanding model (ELU) (198), degraded signals in adverse listening environments force the listener to engage a range of cognitive processes for top-down strengthening of the target signal (199, 200). The most prominent cognitive changes in MS are slowed cognitive processing speed, attention, memory (episodic) and visuospatial skills, with additional impairments in executive function and verbal fluency (57-59). Consequently, the

use of multifaceted sensory-and-cognitive task of SiN processing tasks may offer a sensitive method for capturing cognitive speech impairments in early, minimally disabled MS.

A SiN task comprises the registration of auditory input, the deployment of central cognitive resources to extract the speech of interest (201), and finally the generation of a verbal response that repeats the targeted speech of interest. A cognitive test commonly used in MS is the Paced Auditory Serial Addition Test (PASAT); a complex test of working memory, mental arithmetic, and information speed, that similarly requires the registration of auditory input, the deployment of central cognitive resources, and the generation of a verbal response (202). At the time the Multiple Sclerosis Functional Composite (MSFC) was developed, the PASAT was included due to its sensitivity to neurocognitive effects of MS, brevity, ease of administration and lack of reliance on visual or motor function (203). However there is increasing encouragement from the field to consider replacing the PASAT with the Symbol Digit Modalities Test (SDMT), a widely used, simple and brief measure of information processing speed (IPS) (204), as there are concerns about whether PASAT deficits are related to the ability to perform the task or to the ability to accurately learn the instruction set for an unfamiliar task (205). Many participants use a 'chunking' strategy which reduces cognitive demand, casting doubt on the reliability of scores (206, 207). There are also persistent complaints that the PASAT is unpleasant and stressful, with one study finding 14.2% of participants were unwilling or failed to complete the test (205), and even healthy controls reacting with aversion (208). It is also worth noting that while simpler tasks that present digits aurally, like the Forward and Backward Digit Span Tests, place a lesser load on working memory, these are insensitive to the subtle changes seen in people with early stage MS (55, 209). This highlights the need for a task with significant cognitive load to demonstrate changes in early stage MS, and we propose that a task with high familiarity would provide greater confidence that participants were completing the task as intended.

Here, we evaluated whether early stages of cognitive decline in pwMS (with normal hearing), can be detected using a SiN discrimination task. Our task is ethologically-relevant, with high familiarity, and so required little pre-training and was highly relevant to everyday life where we routinely process speech in backgrounds of noise. We propose that our SiN task engages a broader set of cognitive processes and places greater cognitive load than a clear speech task. Using sentences (compared to single words) and modulated noise, requires accessing stored lexical knowledge and integrating it with new, partially degraded information to improve comprehension (200, 210). This relies heavily on working memory (211). Here, we employed complete speech sounds of words or sentences over a wide range of signal-to-noise ratios (SNRs), modelling conditions of high clarity through to near incomprehension. We also employed two different background types; one that causes "energetic masking" to diminish audibility of a target from interference of shared spectro-temporal acoustic signals in the lower levels of the auditory system (210), and one involving energetic interference but also "informational masking" that produces high-level attention competition effects due to confusability of similar target and masker (201, 212, 213). As we evaluated whether early stages of cognitive decline in pwMS, we focused primarily on performance in individuals at the early stages of the disease and only pwMS < 5 years after diagnosis/presentation and EDSS < 3 were evaluated. However, individuals

with late mild MS (> 10 years after onset, EDSS <3) were also evaluated as physical and cognitive deficits may develop separately over the course of MS (214). Given that cognitive impairments and central auditory processing deficits are reported in early and mildly impaired MS (215), we hypothesized that pwMS would exhibit deficits in the dynamic auditory and cognitive processes underlying SiN discrimination, and that these deficits would predate overt physical disability.

3.2 Methods

All procedures were approved by the Monash University Human Research Ethics Committee (8170) and conformed to the guidelines of the National Health and Medical Research Council of Australia and the protocols of the Helsinki Declaration for experiments involving human participants.

3.2.1 Participants

MS participants were recruited through Royal Melbourne Hospital Australia, and neurologically healthy controls were recruited from the local community. Only patients with relapsing-remitting or CIS were included here; secondary and primary progressive types were excluded. Relapsing-remitting MS patients were defined based on McDonald's criteria (29) and CIS inclusion was based on the initial neurological disturbance (with varying presentations including visual disturbances, numbness/weakness, and balance problems) and Magnetic Resonance Imaging (MRI) evidence of demyelination. All MS participants were independently mobile, with little to no disability (EDSS of < 3) and continued to take all prescribed medication. No patient experienced exacerbated symptomology for at least 3 months prior to participation.

Exclusion criteria for both MS and control participants were a history of another neurological disorder, substance abuse/dependence, pregnancy, and/or the presence of hearing loss (see section 2.3 Audiometry). All participants reported English as their native language.

3.2.2 Neuropsychological testing

To verify that SiN performance was associated with cognition abilities in MS, neuropsychological testing was conducted in the pwMS. Beck's Depression Inventory (BDI) (216), a self-rating inventory of depression, was used as the presence of depressive symptoms are known to be associated with cognitive performance in MS (217). Total scores between 1- 10 are considered normal; 11-16 a mild mood disturbance; and any score over 31 suggests severe/extreme depression. The National Adult Reading Test (NART), a test of premorbid intellectual functioning, was used to measure cognitive reserve (218). The NART consists of 50 words with atypical phonemic pronunciation, and participants are required to read each aloud (untimed). Higher scores indicate greater cognitive reserve. A modified form of the Fatigue Impact Scale (MFIS) (219) was used for self-reported fatigue on three subscales: physical, cognitive, and psychosocial. Higher total MFIS scores indicate a greater impact of fatigue on a person's activities. Neuropsychological tests evaluated in pwMS included: the PASAT (220), SDMT (221), California Verbal Learning Test (CVLT) (222), and Digit Span Test (DST – WAIS-IV administration) (forward and back).

3.2.3 Pure-tone audiometry

Hearing sensitivity was determined using a Beltone Model 110 Clinical Audiometer and calibrated TDH headphones to test sensitivity one ear at a time, at standard audiometric frequencies of 250Hz, 500Hz, 750Hz, 1000Hz, 1500Hz, 2000Hz, 4000Hz, 6000Hz and 8000Hz, using a modified Hughson-Westlake procedure(174). Hearing thresholds,

recorded as decibels Hearing Level (dB HL) relative to normal sensitivity (ISO 8253-1, 1989), were defined as the lowest level at which the tone was perceived 50% of the time. Pure tone averages of the hearing thresholds levels at 500, 1000, 2000 and 4000 Hz were obtained for all participants.

3.2.4 Auditory attention and discomfort questionnaire (AADQ)

The AADQ was developed by Dunlop, Enticott and Rajan (2016) and based on validated inventories for specific adult clinical populations with abnormal auditory processing(141). The 33-item AADQ consists of statements about daily life events involving hearing and had three subscales; the Audio-Attentional Difficulty subscale measures difficulties attending to speech in noisy environments; the Auditory Discomfort (Non-Verbal) subscale measures discomfort to non-verbal environmental sounds; and the Auditory Discomfort (Verbal) subscale measures discomfort to verbal sounds. Refer to **Supplementary Figure A.3** for details on the questionnaire.

3.2.5 Loudness sensitivity test

Participants were asked to describe the extent of his/her auditory discomfort in response to stimuli presented through headphones. This test of hypersensitivity to sounds was previously conducted in the case of participants with high-functioning Autism Spectrum Disorder (ASD) to interpret difficulties in speech discrimination (141). We employed a procedure similar to that described by Dunlop et al., (2016). Stimuli presented during this test were played manually from .wav files using an in-house program.

A chart was placed in front of the participant with the numbers 1 to 7 drawn in a hemi-circle. Emoticons were placed at the numbers 1, 4 and 7: a smiley face at 1 to indicate no discomfort, a neutral face at 4 to indicate moderate discomfort and a sad face at 7 to indicate great discomfort. Seven sets of three sentences were presented binaurally at levels ranging from 60 dBA (A-weighted decibels) to 90 dBA in 5 dB steps. Sentences were derived from a standard clinically-used battery of sentences, the Bamford-Kowal-Bench (BKB) sentence lists(176) consisting of simple sentences in common use (176).

Multi-talker babble (BN) was also presented binaurally at seven different levels ranging from 65 dBA to 77 dBA in 2 dB steps. BN consisted of eight simultaneous voices generated by doubling over and temporally offsetting a recording of four people reading nonsense text. Participants indicated the extent of auditory discomfort they experienced for each stimulus by pointing to the number that corresponded to their perceived loudness discomfort level. Note: the same BN was used in the speech in discrimination tasks.

3.2.6 Speech in noise discrimination tasks

All participants participated in two tasks involving speech discrimination in noise presented at various SNRs. Speech stimuli in the first task consisted of sentences (SiN), whilst in the second task they consisted of singular words (WiN). Sentences were presented in speech-weighted noise (SWN) and BN whilst singular words were only presented in BN.

The general procedures and stimuli for the SiN task have been detailed previously (141, 175). In brief, speech stimuli were derived from a standard clinically-used battery of sentences, the Bamford-Kowal-Bench (BKB) sentence lists(176) consisting of simple sentences in common use (176), with each containing 4-6 words with 3 being keywords for scoring. Sentences were presented in a masker of SWN or BN. SWN was shaped to the long-term average spectrum of the target sentences(141). BN consisted of eight simultaneous voices generated by doubling over and temporally offsetting a recording of four people reading nonsense text.

Speech and masker stimuli were presented binaurally through Sennheiser HD535 headphones. Sentences were presented at a constant level of 70 dBA, with the masker level varied to generate SNRs of 3, 1, -1, -3, and -5 dB in BN; and 1,-1,-3,-5,and -7 dB in SWN. Prior to each noise condition, participants completed 10 practice trials (10 unique target sentences) at an 'easy' SNR of +5 dB for acclimatization to stimuli and task. Subsequent SNR blocks were presented in random order. At each SNR, 10 unique sentences were presented one at a time and the listener asked to repeat each sentence or indicate inability to do so before the next sentence. Discrimination accuracy in the task was scored by: (1) tallying correctly discriminated sentences (/10 per SNR; all three keywords had to be correct) and (2) separately tallying correctly discriminated keywords within the sentences (/30 per SNR) across all sentences regardless of whether the sentence had been correctly discriminated (i.e., whether all 3 keywords in the sentence had been detected). No time limit was placed on response and feedback was not provided. The experimenter recorded the responses and presented the next sentence after 1.5 second delay.

For the words-in-noise (WiN) task, single keywords from the BKB sentences were used as speech stimuli, and presented in BN. To ensure the words were presented identically acoustically to how keywords were presented in the sentences in the SiN task, the words were sliced carefully from the pre-recorded sentences. Test SNRs were 3, 1, -1, -3, and -5 dB, with tests being preceded by a practice session with a unique set of 30 words (30 trials) at an SNR of 5 dB. For each test SNR too, 30 target words were presented; order of test SNR was randomized after the participant had completed the practice session.

3.2.7 Statistical analyses

Statistical analyses were conducted with IBM SPSS Statistics 26, MATLAB 2019b and GraphPad Prism 8 programs.

Participant demographics and hearing sensitivity were compared across control, early and late mild MS groups by Chi-squared tests, Kruskal-Wallis Tests and One-Way ANOVAs, depending on the distribution of data sets.

Depression, fatigue, premorbid intelligence levels and neuropsychological evaluations were compared between early and late mild MS groups by Mann-Whitney or unpaired Student's t-tests, again depending on the distribution of data sets.

Pure-tone hearing thresholds and all SiN tasks were evaluated using two-way mixed-effects analysis of variance (ANOVA) and post hoc Tukey's multiple comparisons tests. Boltzmann sigmoidal functions were fitted to

obtain psychometric curves as a function of SNR for individual participants in each SiN task. Slope and midpoint data from the curves were compared using one-way ANOVAs.

Pearson's correlations were used to determine the relationship between SiN measures and several clinical and neuropsychological measures; with the exception of the association between EDSS scores and SiN measures, which was run as a Spearman correlation.

The midpoints of the psychometric curve for each SiN task were used in analyses of receiver operating characteristic curves to classify between controls and all pwMS. Areas under the curves (AUC-ROC) were obtained to evaluate classification performance. Youden's Index was used to determine a cut-off point, and sensitivity and specificity were obtained.

A logistic regression model was developed to discriminate between controls (coded as 0) and all low impaired MS participants (coded as 1). The model building strategy was to only consider speech discriminated at certain SNRs as predictor variables. The model was validated using five-fold cross validation, AUC-ROC and confusion matrix.

3.3 Results

3.3.1 Participant demographics, characteristics, and audiometric hearing status

In total 50 controls and 40 pwMS were recruited for this study, however, 12 controls (24%) and 5 pwMS (12.5%) were excluded for bilateral hearing loss (pure tone thresholds ≥ 25dB HL) at some/all frequencies between 250 – 4000 Hz. The remaining 38 controls and 35 pwMS who contributed the data reported here (**Table 1**) had bilaterally normal hearing between 250 – 4000 Hz; of these, 5% from each group had small hearing losses (of 5–10dB) at the higher test frequencies of 6000 and 8000 Hz in one ear only. Of the 35 pwMS participants, 15 relapsing-remitting and 8 CIS participants were classified as early mild-MS (≤ 5 years after diagnosis; EDSS < 3) and 12 relapsing-remitting as late mild-MS (≥ 10 years after diagnosis; EDSS < 3).

Demographics such as age, sex, and mean pure-tone averages (dB HL) did not differ between controls, early mild MS, and late mild MS groups ($p>0.05$). **Figure 1** shows mean pure tone air-conduction thresholds at audiometric test frequencies for left **(A)** and right **(B)** ears of the three groups. A two-way mixed-effects analysis of variance (ANOVA) confirmed that hearing sensitivity was similar for the three groups for both left [$F(2,70) = 0.29$, $p = 0.75$, $\eta^2 = 0.26$] and right ears [$F(2,70) = 0.39$, $p = .68$, $\eta^2 = 0.34$]. In summary, the three groups all had normal hearing that was similar across groups.

With respect to the two MS sub-groups, early mild-MS and late mild-MS, disease duration was statistically significant ($p<0.0001$) but EDSS scores and the percentage of participants on disease modifying therapies were not statistically significant. There were also no differences in performance on the neuropsychological and cognitive test results. Estimated premorbid intelligence, depression, and fatigue (as measured by the NART, BDI and MFIS scores, respectively) were not statistically significant between the early and late mild-MS groups. In particular, performance on tests involving auditory input, such as the PASAT, SDMT, CVLT and digit span tests (forward and backward), also did not differ statistically between early and late mild-MS groups. Thus, although the early mild-MS and late mild-MS sub-groups differed in disease duration, they did not differ on EDSS scores and neuropsychological tests, including tests with an auditory processing component.

Table 1. Participant demographics, disease characteristics, estimated premorbid intelligence, depression, fatigue, and neuropsychological test details

		Control	All MS	Early MS	Late MS	p value
Demographics	**Number of participants**	38	35	23	12	
	Sex F(M)	35 (3)	31(4)	20(3)	11(1)	
	Age, (yrs)					
	Mean (SD)	45.66 (10.43)	44.94 (10.59)	42.86 (11.10)	49.25 (8.32)	0.22 [a]
	Range	28 - 60	26 - 65	26 - 65	37 - 65	
Auditory evaluation	**Pure tone average (dB HL)**					
	Left (Mean, SD)	13.03(4.80)	12.79(4.34)	12.78(4.68)	12.82(3.81)	0.98 [b]
	Right (Mean, SD)	11.97(4.49)	13.11(5.08)	13.26(4.98)	12.81(5.49)	0.58 [b]
	Auditory Attention and Distress Questionnaire					
	Audio-attentional difficulty (Mean, SD) Total/98*	23.63 (8.11)	25.86 (11.13)	24.61(10.33)	28.25(12.63)	0.36 [b]
	Auditory Discomfort (non-verbal)(Mean, SD) Total/56*	26.05 (9.29)	24.06 (8.27)	23.30 (8.59)	25.50 (7.76)	0.50 [b]
	Auditory Discomfort (verbal)(Mean,SD) Total/35*	16.79 (6.34)	17.80 (8.05)	16.61 (6.80)	20.08 (9.95)	0.34 [b]
Disease characteristics	**Disease duration (yrs)**					
	Mean (SD)	NA	7.26 (6.25)	3.14 (1.59)	14.8 (3.63)	**<0.0001** [c]
	Range	NA	0 - 22	0 - 5	10 - 22	
	EDSS~					
	Mean (SD)	-	0.37(0.81)	0.35 (0.82)	0.42 (0.82)	0.69 [c]
	Range	-	0 - 2.5	0 – 2.5	0 – 2.5	
	Phenotype RR(CIS)	-	27(8)	15(8)	12(0)	

	On disease modifying therapy (n, %)	NA	29 (82.9%)	19 (82.6%)	10 (83.3%)	0.99 [d]
Estimated Premorbid Intelligence	**National Adult Reading Test**					
	Mean (SD)	-	116(5.31)	114.9(5.18)	117(5.61)	0.36 [e]
	Range	-	105 - 125	105 - 124	105 - 125	
	Data missing (n,%)	-	9 (25.7%)	6 (26.1%)	3 (25%)	
Depression	**Beck's Depression Index**					
	Mean (SD)	-	4.79(4.43)	4.9(4.32)	4.6(4.9)	0.85 [e]
	Range	-	0-14	0 - 14	0-13	
	Data missing (n,%)	-	6 (17.1%)	4 (17.4%)	2 (16.7%)	
Fatigue	**Modified Fatigue Impact Scale**					
	Mean (SD)	-	26.25(16.08)	28.58(15.87)	21.33(16.31)	0.15 [c]
	Range	-	0 – 49	0 – 49	0 – 39	
	Data missing (n,%)	-	7 (20%)	4 (17.4%)	3 (25%)	
Neuropsychological assessments (delivered in the auditory domain)	**Paced Auditory Serial Addition Test**					
	Mean % (SD)	-	82.83(19.62)	85.87(16.45)	77.03(24.48)	0.35 [c]
	Range	-	28.33 – 100	51.67 - 100	28.33-100	
	Data missing (n,%)	-	6 (17.1%)	4 (17.4%)	2 (16.7%)	
	Symbol Digit Modalities Test					
	Mean (SD)	-	65.66(14.30)	67.68(13.2)	61.8(16.3)	0.40 [e]
	Range	-	31 - 92	49 - 92	31 - 88	

Data missing (n,%)	-	6 (17.1%)	4 (17.4%)	2(16.7%)	
California Verbal Learning Test					
Mean (SD)	-	45.44(14.03)	42.47(13.7)	50.5(13.79)	0.15 [e]
Range	-	10 - 67	10 - 62	19 - 67	
Data missing (n,%)	-	6 (17.1%)	4 (17.4%)	2 (16.7%)	
Digit Span (Forward)					
Mean (SD)	-	11.04(2.68)	11.44(2.78)	10.12(2.36)	0.25 [e]
Range	-	6 – 16	7 - 16	6 – 14	
Data missing (n,%)	-	9 (25.7%)	5 (21.7%)	4 (33.3%)	
Digit Span (Backward)					
Mean (SD)	-	7.27(2.19)	7.5(2.12)	6.75(2.49)	0.44 [e]
Range	-	3 – 11	4 - 11	3 - 10	
Data missing (n,%)	-	7 (20%)	5 (21.7%)	2 (16.7%)	

F = female; M = male

dB HL = decibels hearing level

RR = Relapsing Remitting

CIS = Clinically Isolated Syndrome

SD = Standard deviation

NA = Not applicable

n = Number

~EDSS = Expanded Disability Status Scale Score determined by a neurologist within 6 months of audiological testing

***** Higher total scores in the AADQ were indicative of greater Audio-attentional difficulty (range 14-98); greater non-verbal discomfort (range 8-56); and greater verbal discomfort (range 5-35);

[a] Kruskal-Wallis Test

[b] One-Way ANOVA

[c] Mann-Whitney Test

[d] Fisher's Test

[e] Unpaired Student's T Test

NOTE: Demographics and audiometry metrics were compared across controls, early and late mild MS. Disease characteristics, estimated premorbid intelligence, fatigue, and neuropsychological assessments were compared between early and late mild MS

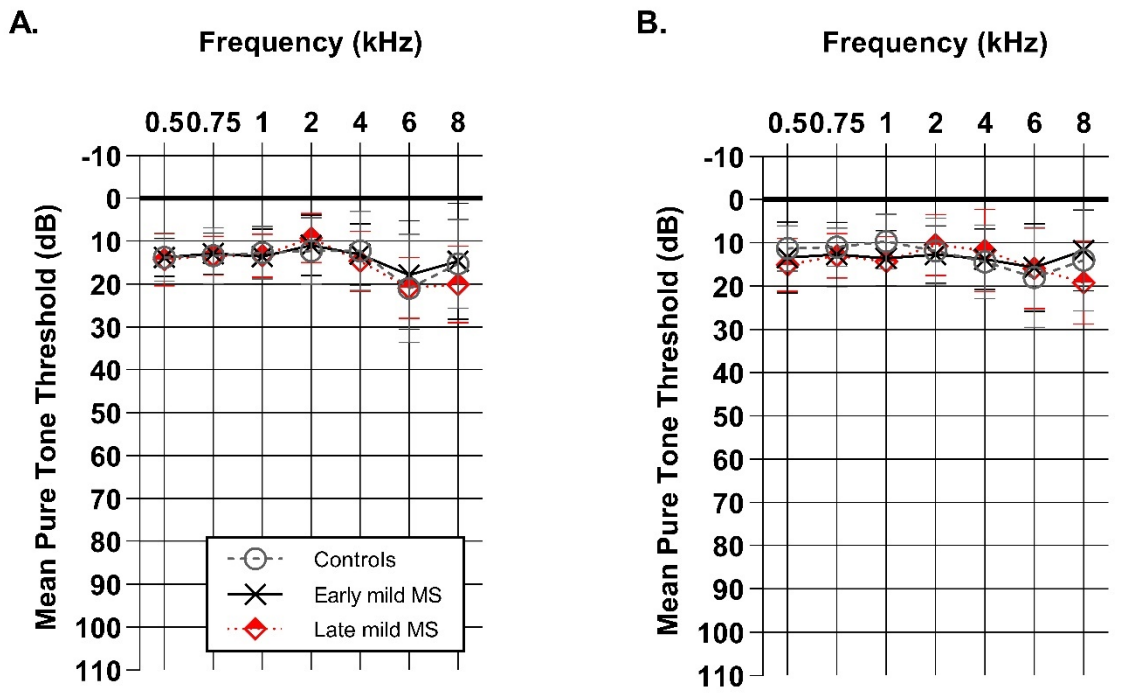


Figure 1. Early and late mild MS groups have similar pure-tone hearing sensitivity (0.5 – 8 kHz) to controls. Mean pure tone thresholds (± standard deviation (SD)) obtained from left **(A)** and right **(B)** ears of control (n = 38; circle/solid line), early mild-MS (n = 23; cross/broken line) and late mild-MS (n = 12; diamond/dotted line) were not significantly different. Two-way mixed ANOVA (p>0.05).

3.3.2 Sensory discomfort to high intensitity stimuli

Mean loudness discomfort levels (LDLs) (±SEM) for speech and BN are represented in **Figure 2A and 2B,** respectively. The trend in both graphs demonstrate that controls and pwMS were, on average, increasingly uncomfortable as speech and noise stimuli were presented at louder intensities (dB) (1= comfortable to 7= discomfort). This trend was confirmed by 3 x 7 two-way mixed ANOVAs which revealed an effect of intensity (dB) on LDLs for speech [$F(6,384) = 115.5$, $p < 0.0001$, $\eta^2 = 39.03$] and noise [$F(6,384) = 106.7$, $p < 0.0001$, $\eta^2 = 25.33$].

An interaction effect between disease group and intensity on speech LDLs was significant [$F(12,384) = 1.89$, $p = 0.03$]. This result and the trend in **Figure 2A** suggest that late pwMS reported higher LDLs than early pwMS and controls only at higher speech intensities. A Tukey's multiple comparisons test revealed that there was no statistical difference at any of the seven intensity levels between disease groups (p>0.05).

There was no significant interaction effect between disease group and intensity on noise LDLs [$F(12,384) = 0.86$, $p = 0.59$], but there was a significant effect of disease group [$F(2,64) = 5.35$, $p = 0.007$]. A Tukey's multiple comparisons test confirmed that late mildly pwMS had significantly higher noise LDLs compared to early mild pwMS ($p = 0.02$) and controls ($p = 0.006$) (**Figure 2B**). All pwMS were tested > 3 months after a relapse of symptoms, hence,

sensitivity to sound was unlikely to be a transient episode of hyperacusis which has been reported in rare case reports (223).

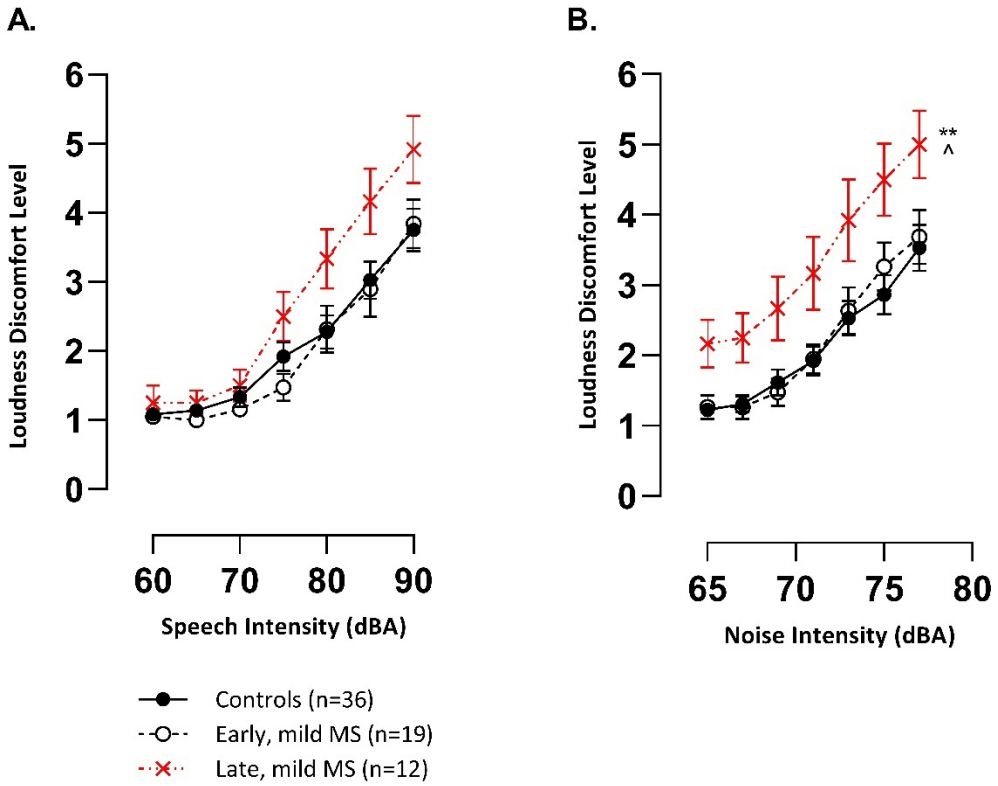


Figure 2. Late mild pwMS have higher discomfort levels to speech and noise (multi-talker babble) stimuli compared to early mild pwMS and controls. Mean Loudness Discomfort Levels (± SEM) of control (n = 36; filled circle/solid line), early mild-MS (n = 19; empty circle/broken line) and late mild-MS (n = 12; cross/dotted line) for speech stimuli **(A)** and multi-talker babble **(B)** presented at various intensities (A-weighted decibels). ** $p < 0.01$ compared to controls; ^ $p < 0.05$ compared to early mild MS. Two-way mixed ANOVA, and Tukey's post hoc test ($p>0.05$).

3.3.3 Discrimination of sentences in noise

In our core tasks, participants were tested with different lists of sentences separately in SWN and in BN. As shown in **Figure 3A&B**, the general trend for all listeners in both conditions was a decrease in sentence identification as SNR decreased. SiN discrimination was easier in SWN than BN as controls were able to correctly identify 50% of the sentences at an SNR of -6.8 ± 0.19 dB in SWN compared to a higher SNR of -0.39 ± 0.38 dB being needed in multi-talker BN for the same level of performance.

3.3.3.1 Discrimination of sentences in speech-weighted noise

Effects in the energetic SWN masker are shown in **Figure 3A.** Sentence discrimination in SWN was relatively easy for SNRs ≥ -1 dB at which controls and mild-MS participants had close to perfect performance recall (98.1%) but at SNRs < -1 dB sentence discrimination degraded for all listeners. These effects were confirmed by a 3 x 5 [i.e., 3 treatment groups (control, early mild-MS, late mild-MS) x 5 SNRs (1, -1, -3, -5, and -7 dB)] two-way mixed ANOVA. No interaction effects were significant between SNR and listener group [$F(8, 280) = 0.70$, $p = .70$, $\eta^2 = 0.27$], but there was a significant main effect of listener group [$F(2,70) = 4.86$, $p = 0.01$, $\eta^2 = 0.85$] and of SNR [$F(4,280) = 328.8$, $p < .0001$, $\eta^2 = 64.02$]. Simple main effects analysis showed that early mild pwMS discriminated fewer sentences than controls ($p = 0.008$), but there were no differences between late mild-MS and controls ($p = 0.24$) or late mild-MS and early mild-MS ($p = 0.43$).

3.3.3.2 Discrimination of sentences in multi-talker babble

Effects in the attentionally-demanding BN are shown in **Figure 3B**. The BN appeared to degrade speech intelligibility for mild-MS listeners more than controls at all SNR conditions except at an SNR of -5 dB at which a floor effect was observed for all groups (**Figure 3B**). A 3 x 5 (i.e., 3 treatment groups x 5 SNRs) two-way mixed ANOVA was used to compare the ability of the groups to discriminate sentences in BN in various SNR conditions. No interaction effects were found [$F(8, 280) = 1.07$, $p = 0.38$, $\eta^2 = 0.23$], however, there was a significant main effect of listener group [$F(2,70) = 6.29$, $p = 0.003$, $\eta^2 = 0.66$] and of SNR [$F(4,280) = 668.8$, $p < .0001$, $\eta^2 = 70.6$]. Simple main effects analysis showed that both early and late mild-MS groups discriminated fewer sentences than controls ($p = 0.03$ and $p = 0.009$, respectively). There was no difference in discrimination between early and late mild-MS groups ($p > 0.99$).

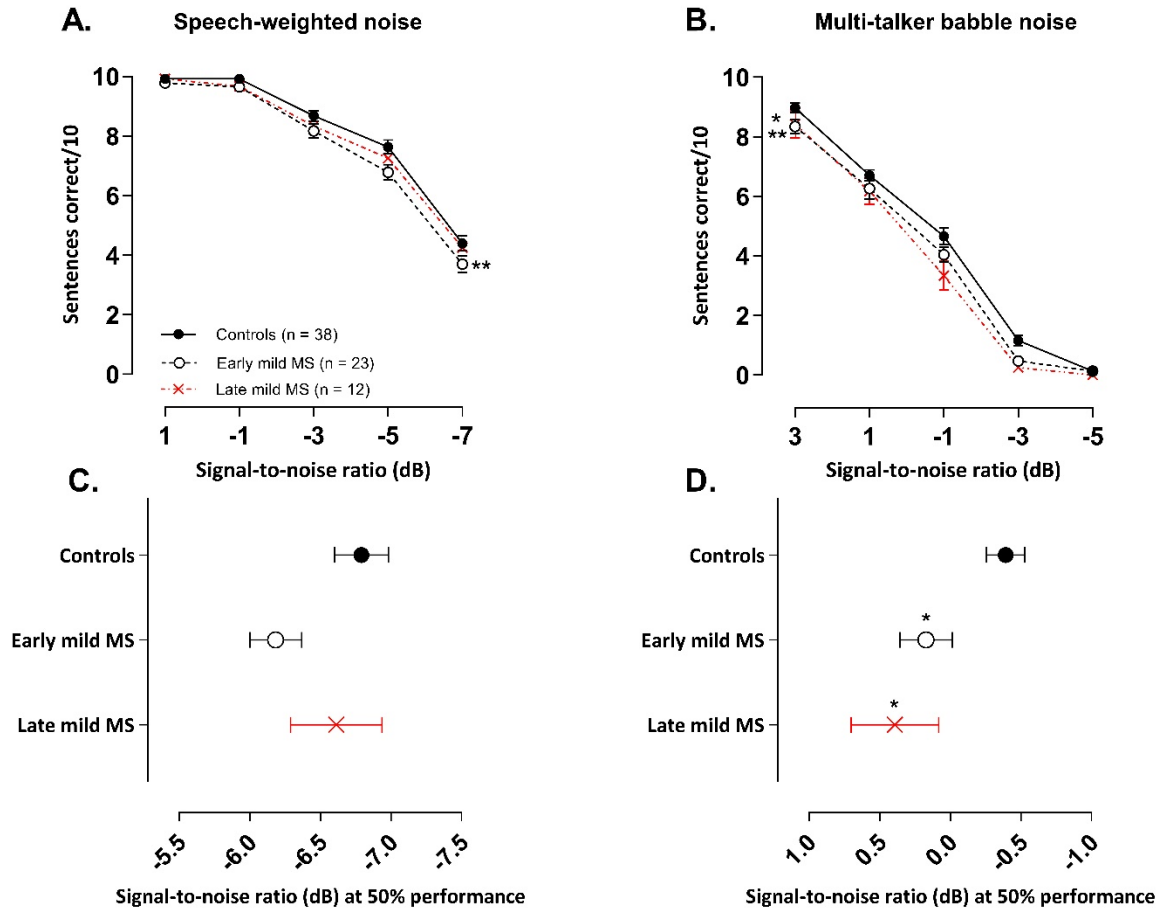


Figure 3. PwMS discriminated fewer sentences than controls in an attention-demanding babble masker, but not in speech-weighted noise. Mean ± SEM correctly discriminated sentences (i.e., all 3 keywords detected in a sentence) out of 10 test sentences at each signal-to-noise ratio (SNR (dB)) for control (n = 38, filled circles/solid line), early (n = 23, open circle/dashed line) and late mild MS (n = 12, cross/dotted line) in speech-weighted noise **(A)** and multi-talker babble **(B).** Mean SNRs ± SEM (dB) at 50% discrimination in speech-weighted noise **(C)** and multi-talker babble **(D)**. *p < 0.05; **p < 0.01; compared to controls. (**A&B**: Mixed-effects two-way ANOVA; **C&D:** One-Way ANOVA; all cases Tukey's post hoc test).

To quantify MS effects on sentence discrimination, Boltzmann sigmoidal functions were fitted to each participant's discrimination curves, with the top and bottom of the functions constrained to 10 and 0 sentences correct, respectively. From each such psychometric curve, the slope and midpoint SNR were extracted (see **Table 2**). Measures of goodness of fit were strong for each group (R^2 always > 0.9). A one-way ANOVA revealed no significant difference in mean slopes of the psychometric functions for control, early and late mild-MS groups for speech discrimination in SWN [$F(2,70) = 1.10$, $p = 0.35$, $\eta^2 = 0.03$] or in BN [$F(2,70) = 0.50$, $p = 0.61$, $\eta^2 = 0.01$].

The midpoint SNRs of the curves are graphed in **Figure 3C** and **3D** for sentence discrimination in SWN and BN, respectively. A one-way ANOVA indicated that the midpoint SNRs of control, early and late mild-MS psychometric functions for sentence discrimination in SWN were not significantly different [$F(2,70) = 2.25$, p =

0.11, $\eta^2 = 0.06$]. In contrast, midpoint SNRs of control, early and late mild-MS psychometric functions for speech discrimination in BN were significantly different [$F(2,70) = 4.9$, $p = 0.01$, $\eta^2 = 0.12$] A Tukey's multiple comparisons test confirmed that early-mild and late-mild MS participants required significantly higher SNRs for 50% discrimination accuracy compared to controls ($p = 0.04$ and 0.02, respectively). There was no significant difference between the SNRs of the midpoints of the curves for the two MS groups ($p = 0.76$).

At the SNR (-0.39 ± 0.13 dB) at which controls attained 50% sentence intelligibility in BN, speech intelligibility in early mild pwMS and late mild pwMS was 9.11 ± 0.21% and 13.96 ± 0.22% lower, respectively.

Table 2. Degrees of freedom (df), goodness of fit (R^2), slope ± SE and midpoint ± SE values for Boltzmann sigmoidal functions fitted to the mean sentences correctly discriminated at tested signal-to-noise-ratios.

		df	R^2	Midpoint ± SE SNR (dB)	Slope ± SE (Sentences/dB)
Speech-weighted noise	**Controls**	37	0.91	-6.79 ± 0.19	1.54 ± 0.13
	Early mild MS	22	0.93	-6.19 ± 0.18	1.78 ± 0.14
	Late mild MS	11	0.92	-6.61 ± 0.32	1.85 ± 0.23
Multi-talker babble noise	**Controls**	37	0.92	-0.39 ± 0.13	1.43 ± 0.08
	Early mild MS	22	0.93	**0.17 ± 0.19***	1.52 ± 0.09
	Late mild MS	11	0.93	**0.39 ± 0.31***	1.36 ± 0.14

df = degrees of freedom
R^2 = goodness of fit
SNR = signal-to-noise ratio
SE = standard error
dB = decibels
(* $p < 0.05$; compared to control in babble, One- Way ANOVA with Tukey's multiple comparisons post hoc test)

3.3.4 Discrimination of keywords in noise

In the SiN tasks detailed above, especially at lower SNRs, listeners were often able to identify some of the words in a sentence but not all three keywords required to score correct discrimination of the whole sentence. We therefore conducted a second analysis where we examined the number of keywords detected correctly across all 10 sentences (30 keywords total at 3/sentence) for each SNR block in the SiN task whether the sentence in which the keyword was embedded was scored correct or not.

3.3.4.1 Discrimination of keywords in speech-weighted noise

The mean number of keywords (± SEM) correctly discriminated by controls, early and late mild MS listeners in each SNR block in the SWN masker is presented in **Figure 4A**. A two-way mixed ANOVA indicated that there was no interaction effect between SNR and listener group [$F(8, 280) = 1.04$, $p = .41$, $\eta^2 = 0.43$], but there was a significant main effect of listener group [$F(2,70) = 4.61$, $p = 0.01$, $\eta^2 = 0.78$] and of SNR [$F(4,280) = 322.9$, $p < .0001$, $\eta^2 = 66.2$]. Simple main effects analysis showed that early mild-MS participants discriminated fewer sentences than controls (p

< 0.05), but there were no differences between late mild-MS and controls or late mild-MS and early mild-MS (p > 0.05).

3.3.4.2 Discrimination of keywords in multi-talker babble

The mean number of keywords (± SEM) correctly discriminated by controls, early and late mild MS listeners in each SNR block in the BN masker is presented in **Figure 4B**. A two-way mixed ANOVA indicated the presence of a significant interaction effect between SNR and listener group [F (8, 280) = 2.22, p = .03, η^2 = 0.31], as well as a significant main effect of listener group [F(2,70) = 11.40, p < 0.0001, η^2 = 1.18] and SNR [F (4,280) = 1031, p < .0001, η^2 = 72.91]. Simple effects analysis showed that early, and late mild-MS participants discriminated fewer sentences than controls at an SNR of -1 dB (p = 0.007 and p = 0.0002, respectively) and at an SNR of -3 dB (p = 0.03 and p = 0.001, respectively). There was no significant difference in discrimination between early and late mild-MS groups at either of these SNRs (p > 0.05).

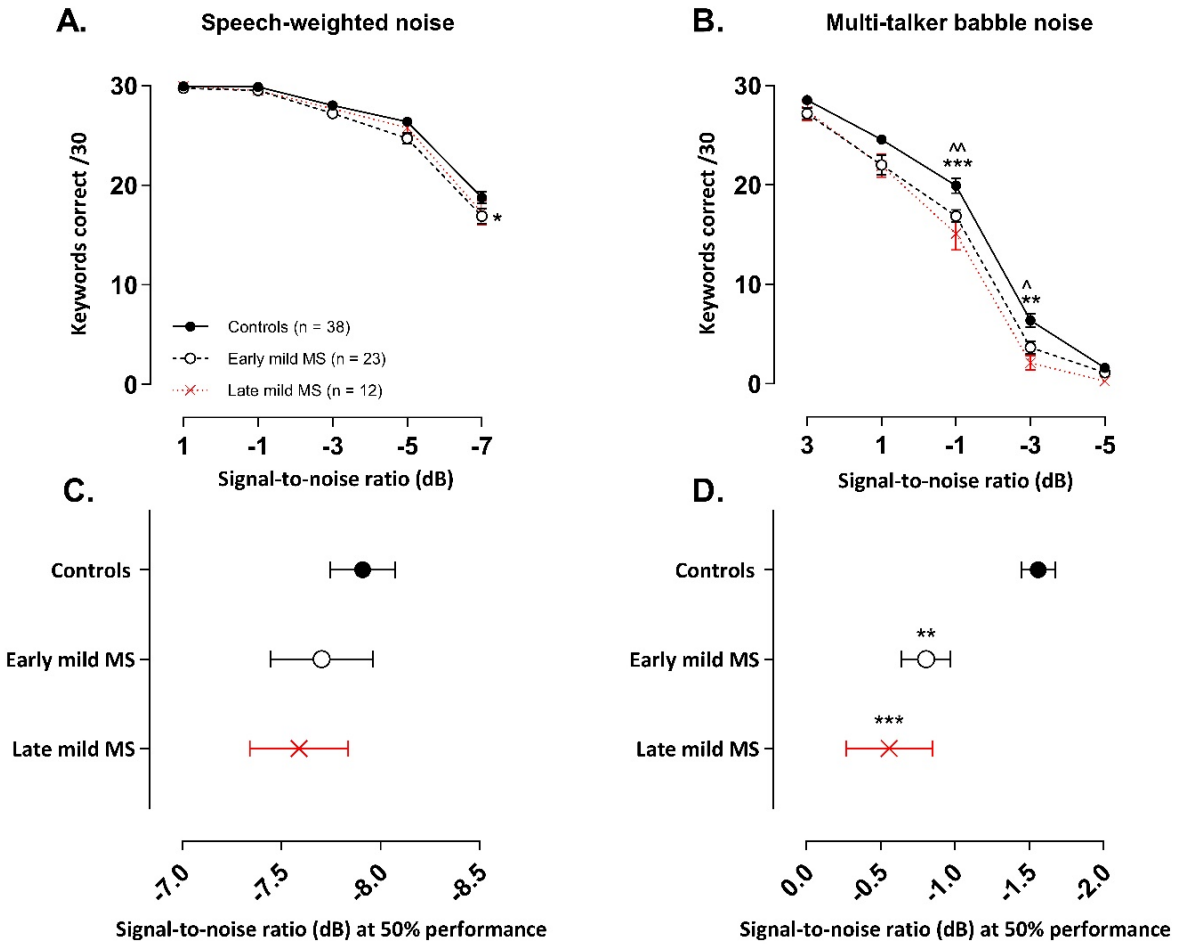


Figure 4. PwMS discriminated fewer keywords than controls in an attention-demanding informational babble masker. Mean ± SEM keywords (/30 at each signal-to-noise ratio (SNR (dB))) correctly discriminated for control (n = 38, filled circles/solid line), early (n = 23, open circle/dashed line) and late mild-MS (n = 12, cross/dotted line) in speech-weighted noise **(A)** and multi-talker babble **(B).** Mean SNRs ± SEM (dB) at 50% discrimination in SWN **(C)** and BN **(D)**. (*)(^)p < 0.05; (**)(^^)p < 0.01; ***p< 0.001; compared to controls. (**A&B**: Mixed-effects two-way ANOVA; **C&D:** One-Way ANOVA; all cases Tukey's post hoc test).

To quantify MS effects on keyword discrimination, Boltzmann sigmoidal functions were fitted to each participant's discrimination curves. The top and bottom of the functions were constrained to 30 and 0 keywords correct respectively. Measures of goodness of fit were strong for each group ($R^2 \geq 0.9$). **Table 3** displays the slope and midpoint data extracted from the psychometric curves.

A one-way ANOVA revealed no significant difference in mean slopes of control, early and late mild-MS psychometric functions for keyword discrimination in SWN [$F(2,69) = 1.27$, $p = 0.29$, $\eta^2 = 0.04$] or BN [$F(2,70) = 3.14$, $p = 0.051$, $\eta^2 = 0.08$]. The midpoints of the curves are graphed in **Figure 4C** and **4D** for discrimination in SWN and BN, respectively. A one-way ANOVA indicated that there was no significant difference in mean midpoints of

control, early and late mild-MS psychometric functions for keyword discrimination in SWN [$F(2,69) = 0.53$, $p = 0.59$, $\eta^2 = 0.02$].

In contrast, there was a significant difference in mean midpoints of control, early and late mild-MS psychometric functions for keyword discrimination in BN [$F(2,70) = 10.84$, $p < 0.0001$, $\eta^2 = 0.24$]. A Tukey's multiple comparisons test indicated that early and late mild-MS participants needed significantly higher SNRs to achieve 50% correct performance compared to controls ($p = 0.002$ and 0.0007, respectively). There was no significant difference between the midpoints of the curves for early and late mild MS participants ($p = 0.65$).

At the same SNR (-1.56 ± 0.11 dB) at which controls attained 50% keyword intelligibility in BN, intelligibility was 13.08 ± 0.54% and 20.00 ± 0.55% lower for early mild pwMS and late mild pwMS, respectively.

Table 3. Degrees of freedom (df), goodness of fit (R^2), slope ± SE and midpoint ± SE values for Boltzmann sigmoidal functions fitted to the mean keywords correctly discriminated at tested signal-to-noise-ratios in speech-weighted and babble noise

		df	R^2	Midpoint ± SE SNR (dB)	Slope ± SE (Keywords/dB)
Speech-weighted noise	**Controls**	37	0.93	-7.91 ± 0.16	1.51 ± 0.11
	Early mild MS	22	0.94	-7.70 ± 0.25	1.79 ± 0.18
	Late mild MS	11	0.95	-7.59 ± 0.25	1.50 ± 0.18
Multi-talker babble noise	**Controls**	37	0.96	-1.56 ± 0.11	1.18 ± 0.06
	Early mild MS	22	0.94	**-0.81 ± 0.16****	1.40 ± 0.10
	Late mild MS	11	0.95	**-0.56 ± 0.29*****	1.18 ± 0.11

df = degrees of freedom
R^2 = goodness of fit
SNR = signal-to-noise ratio
SE = standard error
dB = decibels
(** $p < 0.01$; *** $p < 0.001$, compared to controls in babble, One-Way ANOVA with Tukey's multiple comparisons post hoc test)

3.3.5 Words in babble discrimination

Sentences provide additional syntactic and semantic cues that the listener can use to infer the meaning of partially masked or degraded speech (224). For this reason, SiN tasks often employ simpler stimuli, like phonemes or isolated words in noise to assess speech discrimination ability; despite the use of sentences reflecting better communication demands in the real listening world. We therefore examined whether the effects seen above with whole sentences or with keywords embedded in sentences would be replicated with isolated words, in a background of BN. For this task, individual keywords were extracted from the pre-recorded sentences of the BKB sentence lists used above, and were

presented individually in random order (thereby removing any linguistic context) at similar SNRs as used in the SiN task with the same BN (3, 1, -1, -3, -5 dB).

Not all participants were able to attend the additional session where this test was conducted and so 20 controls, 15 early, and 10 late mild MS participants completed the WiN task. The means and standard errors of correctly recalled words in BN for controls and MS participants are presented in **Figure 5A.** A 3 x 5 two-way mixed ANOVA revealed main group effects were significant [$F(2,49) = 26.96$, $p < .0001$, $\eta^2 = 3.10$]. Additionally, as expected, a decrease in SNR significantly negatively impacted word discrimination in all listener types [$F(4,196) = 656.0$, $p < .0001$, $\eta^2 = 76.2$]. Decreasing SNR degraded speech discrimination performance similarly for all listeners, as evident by no interaction effect [$F(8, 196) = 1.50$, $p = 0.16$, $\eta^2 = 0.35$]. Simple main effects analysis showed that both early and late mild-MS participants discriminated fewer sentences than controls ($p < 0.0001$). There was no difference in discrimination between early and late mild-MS groups ($p = 0.50$).

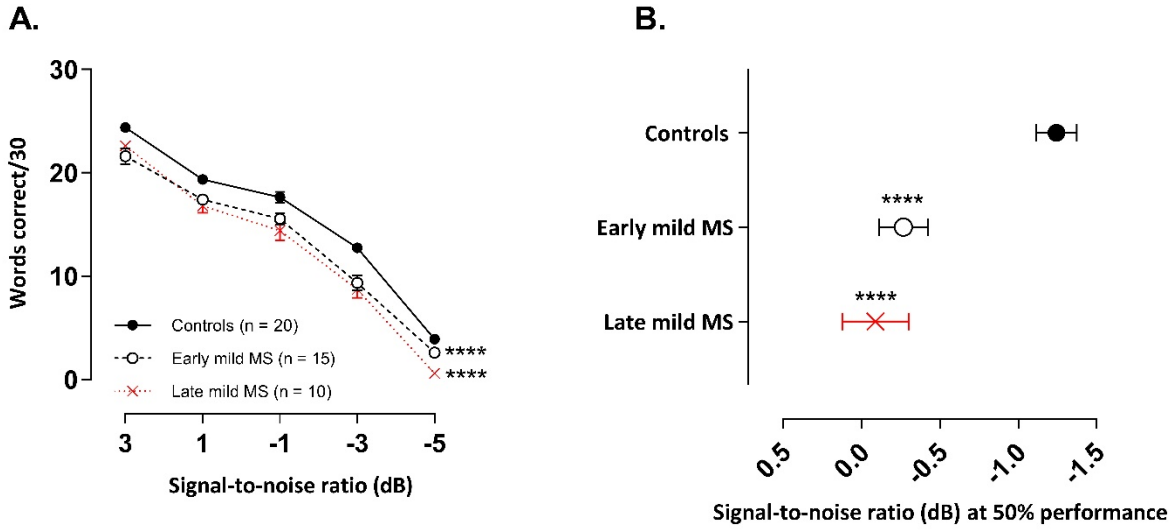


Figure 5. PwMS discriminated fewer words in multi-talker babble compared to healthy controls. Mean ± SEM words (/30 at each signal-to-noise ratio (SNR)) correctly discriminated for control (n = 20, filled circles/solid line), early (n = 15, open circle/dashed line) and late mild MS (n = 10, cross/dotted line) in multi-talker babble **(A).** Mean SNRs ± SEM (dB) at 50% discrimination in BN **(B)**. ****p < 0.0001; compared to controls. (**A** Mixed-effects two-way ANOVA; **B** One-Way ANOVA; all cases Tukey's post hoc test).

Again, to quantify MS effects on word discrimination, Boltzmann sigmoidal functions were fitted to each participant's discrimination curves. **Table 4** displays the slope and midpoint data extracted from the psychometric curves. The top and bottom of the functions were constrained to 30 and 0 words correct, respectively. Measures of goodness of fit were strong for each group ($R^2 \geq 0.88$). A one-way ANOVA confirmed no significant difference in mean slopes of psychometric functions for control, early and late mild-MS groups for word discrimination in BN [$F(2,42) = 3.03$, $p = 0.06$, $\eta^2 = 0.13$]. In contrast, there was a significant difference in mean psychometric function midpoints [$F(2,42) = 16.55$, $p < 0.0001$, $\eta^2 = 0.44$]. The midpoints of the curves are graphed in **Figure 5A**. A Tukey's

multiple comparisons test indicated that the midpoints of the curves for early and late mild-MS participants were at significantly higher SNRS compared to those for controls ($p < 0.0001$). There was no significant difference between the midpoints of the curves of early and late mild-MS participants($p = 0.76$). Thus, these data showed that the psychometric functions for the early and late mild-MS groups had shifted to the left.

At the same SNR (-1.24 ± 0.13 dB) at which controls attained 50%-word intelligibility in BN, intelligibility was 7.86 ± 0.32% and 11.17 ± 0.38% lower in early mild pwMS and late mild pwMS, respectively.

Table 4. Degrees of freedom (df), goodness of fit (R^2), slope ± SE and midpoint ± SE values for Boltzmann sigmoidal functions fitted to the mean words correctly discriminated at tested signal-to-noise-ratios in babble noise

	df	R^2	Midpoint ± SE SNR (dB)	Slope ± SE Words/dB
Controls	19	0.90	-1.24 ± 0.13	2.91 ± 0.10
Early mild MS	14	0.88	**-0.27 ± 0.16****	3.06 ± 0.19
Late mild MS	9	0.88	**-0.09 ± 0.21****	2.53 ± 0.07

df = degrees of freedom
R^2 = goodness of fit
SNR = signal-to-noise ratio
SE = standard error
dB = decibels
(**** $p < 0.001$, compared to controls in babble, One-Way ANOVA with Tukey's multiple comparisons post hoc test)

Across the three discrimination in BN tasks (sentences, keywords, and word discrimination), the similarity of slopes across all three groups showed that any changes across groups were not in the shape of the curves but in the curve location along the SNR axis. This was confirmed by the differences in curve midpoints; i.e., mild MS participants needed more favourable SNRs to achieve the same level of performance. The effect sizes (η^2) for differences in curve midpoints for sentence, keyword and word discrimination in BN were 0.12, 0.24, 0.44, respectively (i.e. 12%, 24% and 44% of the total variance accounted for by the group). Comparison of the effect sizes showed that this shift was most pronounced for the isolated words (absent contextual cues), less so for the keywords embedded in sentences, and least when full contextual and semantic cues were present in the whole sentences.

Our analyses thus far have shown clearly that although early and mild late-MS participants differed significantly with regard to duration of disease, they did not differ significantly on any other metric of the disease or any of the speech performance measures or the neuropsychological tests. Thus, for these groups of pwMS, the determining characteristic appeared to be the fact that they had mild MS. Hence for all subsequent analyses, we pooled these two sub-groups of mild-MS participants into a single pool of people with mild MS.

3.3.6 Correlations of speech discrimination measures to standard neuropsychological tests

Most of our pwMS group underwent neuropsychological testing according to standardized instructions (Refer to **Table 1** for missing data details). We compared performance on these tests against SiN performance, indexing the latter using the midpoints of the psychometric curves (i.e. SNR at 50% speech intelligibility) since that metric had differed significantly from control values (for sentence or word tasks in BN) whereas the slopes of the psychometric functions had not. A Pearson product-moment correlation coefficient was computed to assess the relationship between performance (out of 100%) in the neuropsychological assessments and the midpoints of the psychometric curves, whilst Spearman correlation coefficients were determined for the relationship between EDSS scores and SiN measures. Correlation coefficients between clinical and discrimination measures are displayed in **Table 5**. There was a significant negative correlation with PASAT, SDMT and CVLT scores and the midpoints of the psychometric functions obtained in the sentences in babble task (poorer performance on the neuropsychological tests related to poorer performance on the SiN task).There was also a significant negative association between PASAT and CVLT scores and the midpoints of the psychometric functions obtained in the keywords in babble task. Only the CVLT correlated with the words in babble task. No significant correlations were found between any clinical measures and the midpoints of the psychometric functions for any speech discrimination tests in SWN. No significant relationships were observed between any of the speech discrimination measures and age, disease duration, EDSS, BDI, NART or digit span tests (forward and backward).

Table 5. Correlation coefficients between neuropsychological/clinical measures and speech-in-noise measures (r, (rs^),95% confidence intervals) in all mildly impaired MS participants.

	SWN		**BN**		
	Sentences	*Keywords*	*Sentences*	*Keywords*	*Words*
Age (yrs)	-0.22 [-0.51 to 0.12]	-0.30 [-0.57 to 0.03]	0.19 [-0.15 to 0.49]	0.07 [-0.26 to 0.39]	-0.17 [-0.52 to 0.24]
Disease Duration (yrs)	-0.19 [-0.49 to 0.16]	0.03 [-0.31 to 0.36]	0.15 [-0.19 to 0.46]	0.13 [-0.21 to 0.45]	0.14 [-0.26 to 0.50]
EDSS^	0.32 [-0.03 to 0.59]	0.00 [-0.33 to 0.33]	0.23 [-0.11 to 0.53]	0.12 [-0.22 to 0.44]	0.12 [-0.3 to 0.50]
BDI	0.00 [-0.36 to 0.36]	0.05 [-0.32 to 0.40]	0.09 [-0.28 to 0.44]	0.07 [-0.30 to 0.42]	0.30 [-0.03 to 0.57]
NART	-0.25 [-0.57 to 0.15]	-0.21 [-0.55 to 0.19]	-0.32 [-0.63 to 0.07]	-0.32 [-0.63 to 0.07]	-0.33 [-0.67 to 0.13]
PASAT	-0.02 [-0.38, 0.34]	-0.28 [-0.58, 0.09]	**-0.58*** [-0.78, -0.27]**	**-0.43* [-0.68, -0.08]**	-0.39 [-0.69, 0.02]
SDMT	-0.05 [-0.41, 0.33]	-0.15 [-0.49, 0.23]	**-0.42* [-0.68, -0.07]**	-0.25 [-0.57, 0.12]	-0.13 [-0.52, 0.29]
CVLT	-0.24 [-0.57 to 0.14]	-0.25 [-0.57 to 0.14]	**-0.55** [-0.77 to -0.21]**	**-0.39* [-0.67 to -0.01]**	**-0.44* [-0.73, -0.01]**
Digit Span (Forward)	-0.15 [-0.51, 0.24]	-0.12 [-0.48, 0.28]	-0.19 [-0.54, 0.21]	-0.23 [-0.56, 0.17]	-0.29 [-0.63, 0.16]
Digit Span (Backward)	-0.13 [-0.48, 0.27]	-0.01 [-0.39, 0.38]	-0.31 [-0.62, 0.08]	-0.38 [-0.66, 0.01]	-0.38 [-0.66, 0.01]

SWN = speech-weighted noise; BN = babble noise; EDSS = Expanded Disability Status Scale; BDI = Beck's Depression Index; NART = National Adult Reading Test; PASAT = Paced Auditory Serial Addition Test; SDMT = Symbol Digit Modalities Test; CVLT = California Verbal Learning Test.

^Correlation values are Pearson correlations coefficients (r), with the exception of the associations between EDSS and SiN measures, which are expressed as Spearman correlation coefficients (rs).
* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

3.3.7 Comparison of all SiN measures to discriminate between controls and all MS group

Receiver operating characteristic curves (ROC) were used to evaluate the ability of SiN measures (SNR at 50% discrimination) to discriminate between healthy controls and all pwMS. **Table 6** shows the number of observations obtained for each task, time to administer the test, the cut off SNR point (Youden's J statistic) at which participants were categorized as control or MS, sensitivity (true positive rate), specificity (true negative rate) and the area under the curve (AUC). Interpretation of the AUC ROC indicates that sentence and keyword discrimination in SWN was not useful in discriminating between control and MS participants. In contrast, keyword discrimination in BN was acceptable (0.7 – 0.8). Word discrimination in babble provided outstanding discrimination ability (>0.9) (225).

Table 6. The utility of SiN tasks in discriminating between controls and all mild MS: details of the Receiver Operating Characteristic curves

	Number of observations	Time (mins) to administer test	Cut-off SNR (dB) at 50% discrimination^	Sensitivity (%)	Specificity (%)	ROC AUC (95% C.I)
1) Sentences in…						
a) SWN	73	10	-6.33	51.43	71.05	0.62 (0.49 – 0.75)
b) BN	73	10	-0.75	94.29	39.47	**0.69 (0.56 – 0.81)***
2) Keywords in…						
a) SWN	73	10	-7.29	48.57	78.38	0.58 (0.45 – 0.72)
b) BN	73	10	-1.43	85.71	63.16	**0.79 (0.69 – 0.89)****
3) Words in….						
a) BN	45	20	-0.83	92.0	85.0	**0.91 (0.82 – 0.99)****

^cutoff calculated as Youden's Index
SWN = speech-weighted noise
BN = babble noise
SNR= signal-to-noise ratio
dB = decibels
ROC AUC = Area under the Receiver Operating Characteristic Curve
CI = Confidence Intervals (95%)
** ($p<0.01$); **** ($p<0.0001$). p values test the null hypothesis that the AUC = 0.5

3.3.8 Classification value of speech discrimination tasks to classify all mild (early and late) pwMS from controls

Quantifying neurological functional deficits in mild MS will help clinicians to maximize the opportunities to preserve neurological reserve in patients with appropriate therapeutic management, particularly in the earliest stages. Given that physical assessments are not informative in this fully ambulatory cohort, a logistic regression model was developed using speech discrimination abilities to classify those without MS (controls = 0), and those with mild MS (coded as 1) with a median EDSS score of 0.

To be clinically viable, assessments should be quick and easy to administer, therefore, part of the model building strategy was to only consider speech discriminated at certain SNRs as predictor variables (2-3 minutes testing at each SNR). Midpoint curve SNRs require the whole psychometric function to be obtained; taking anywhere between 10 to 20 minutes to administer the test (refer to **Table 6**); compromising tolerability of the test. Based on the two-way ANOVAs described in **Figures 3,4&5,** performances at specific SNRs in all three BN tasks (sentences, keywords, and words) were considered. Twenty possible regression models (**Table 7**) were generated using MATLAB Statistic Toolbox Release 2019b and based on the SNRs of 3, 1, -1 and -3 dB. -5 dB was not considered as it produced floor effects in the sentences and keywords task. A maximum of two fixed effects were used in any model to avoid overfitting the data with a small data set (45 observations). The model with the lowest AIC was used to select the final model.

The model with the lowest AIC value had two fixed effects: word discrimination in BN (total/30) and sentence discrimination in BN (total/10) at an SNR of -3 dB (model 15 in bold in **Table 7**). Variables had a variance inflation factor (VIF) of 1.23. This was well below the recommended cut off VIF of 5, indicating no problematic levels of multicollinearity among predictors.

Table 7 AIC comparisons of logistics regression models used to classify controls from mild pwMS

Variables in model (with intercept)	Akaike's Information Criterion (AIC)
Variables at SNR of 1dB:	
(1) Words at SNR 1dB	53.69
(2) Keywords at SNR 1dB	62.76
(3) Sentences at SNR 1dB	65.81
(4) Words at SNR 1dB and Keywords at SNR 1 dB	54.89
(5) Words at SNR 1dB and Sentences at SNR 1 dB	55.34
Variables at SNR of -1dB:	
(6) Words at SNR -1dB	56.82
(7) Keywords at SNR -1dB	54.07
(8) Sentences at SNR -1dB	59.69
(9) Words at SNR -1dB and Keywords at SNR -1 dB	51.31
(10) Words at SNR -1dB and Sentences at SNR -1 dB	55.73
Variables at SNR of -3dB:	
(11) Words at SNR -3dB	50.06
(12) Keywords at SNR -3dB	51.22
(13) Sentences at SNR -3dB	52.33
(14) Words at SNR -3 dB and Keywords at SNR -3 dB	39.69
(15) Words at SNR -3dB and Sentences at SNR -3 dB	**38.89**
Variables at SNR of 3dB:	
(16) Words at SNR 3dB	58.58
(17) Keywords at SNR 3dB	56.65
(18) Sentences at SNR 3dB	57.74
(19) Words at SNR 3dB and Keywords at SNR 3dB	55.52
(20) Words at SNR 3dB and Sentences at SNR 3dB	56.05

AIC = Akaike's Information Criterion score is a comparative value that evaluates goodness-of-fit between models. The model with the lowest AIC indicates a superior balance between goodness of fit and avoiding overfitting the data

Table 8 presents the results from a log likelihood ratio test to ascertain if the model with the SiN predictors was more effective than a null model (intercept only). The results of the test suggest that the null model should be rejected in favour of the logistic regression model using SiN measures as predictors $X^2(2) = 18.8$, $p < 0.0001$.

The parameter estimates of fixed effects are also listed in **Table 8**, along with the standard error, t statistic, p values, odds ratio (OR) and CIs (95%). Statistical significance of individual regression coefficients (βs) were tested using the t-statistic (testing the null hypothesis that β is equal to zero). Total sentences and words correctly

discriminated in BN at an SNR of -3 dB were significant discriminators of controls from MS participants (p = 0.004). When all other factors are held constant, for a one unit increase in sentence discrimination performance, the expected OR of the participant being a pwMS was .18 (95% CI: .05 – .60); i.e. 82% (1-$e^{-1.71}$) reduced odds of the participant being a pwMS. There was also an effect of word discrimination, where, for a one unit increase in word discrimination performance, the expected OR of the participant being a pwMS was .48 (95% CI: .29 – .80); i.e. 52% (1-$e^{-0.73}$) reduced odds of the participant being a pwMS.

The logistic regression model was evaluated using 5-fold cross validation, split into a 70:30 training/test set. A receiver operating characteristic (ROC) curve is visualized in **Figure 6** to evaluate the logistic regression as a discrimination tool. The AUC-ROC was 0.86, considered to be excellent classifying performance(225). The cutoff point (Youden's J statistic) was 0.68, and this was used to classify participants into controls and MS. The predicted vs. observed classifications are presented as a confusion matrix in **Table 9**. The model has 80% sensitivity and 85% specificity in the classification of participants. The classification model is not suggested to be used as a diagnostic tool, but as a way to distinguish controls with normal neurological functioning from people with mild MS with subtle neurological dysfunction that goes undetected by the EDSS. Given that the logistic regression has 80% sensitivity in detecting subtle MS deficits, the confusion matrix provides evidence for the 5-minute SiN predictors to be a useful clinical tool.

Table 8. Logistic Regression Model used to classify controls (n= 20) and all mild MS (n=25)

						95% C.I for e^{β}	
Predictor	**Estimate (β)**	**SE β**	**tStat**	***p***	**OR (e^{β})**	**Lower**	**Upper**
Intercept	9.76	3.20	3.05	0.002			
Sentences presented in babble at SNR of -3 dB	-1.71	0.59	-2.88	0.004	0.18	0.05	0.60
Words presented in babble at SNR of -3 dB	-0.73	0.25	-2.91	0.004	0.48	0.29	0.80
		X^2	***df***	***p***			
Overall model evaluation							
Likelihood ratio test		29	2	< .0001			

SNR = signal-to-noise ratio
SE = standard error
OR = odds ratio
C.I = confidence interval
Note: MATLAB 2019b statistical package was used.
Ordinary R^2= 0.53. Adjusted R^2 = 0.51.
45 observations, 42 error degrees of freedom

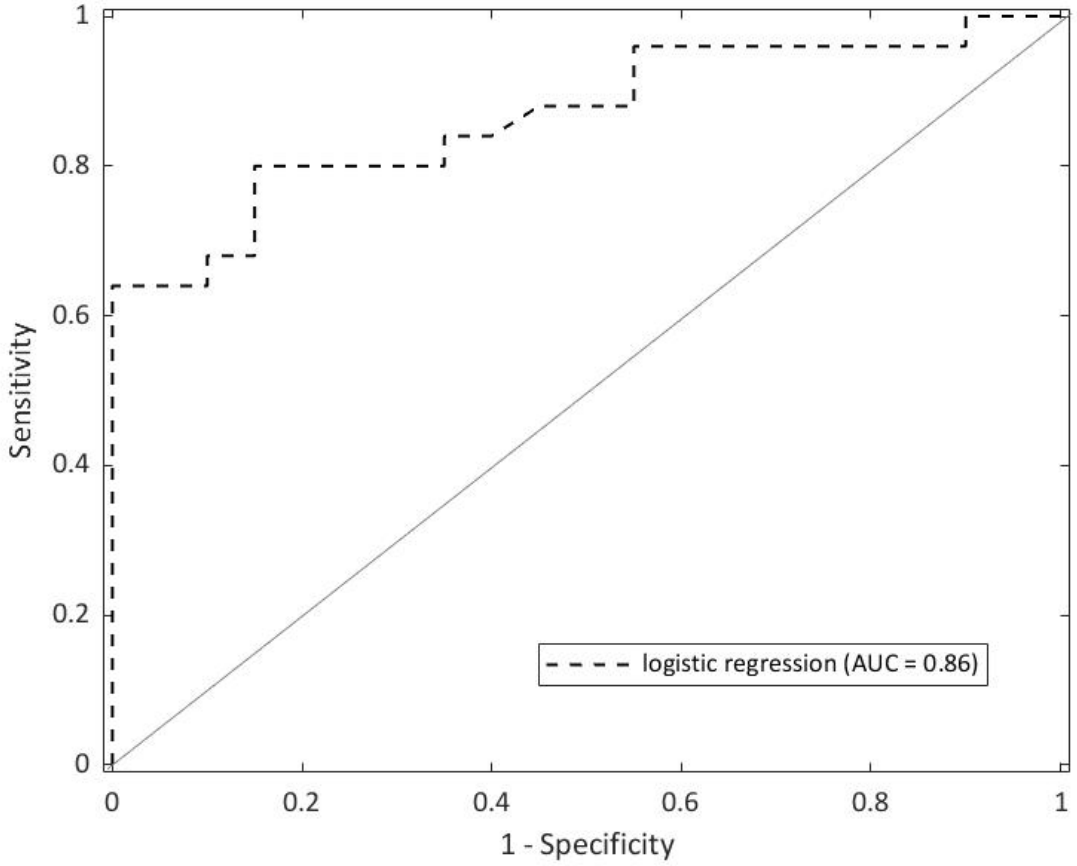


Figure 6. Receiver operating curve for a logistic regression model with predictor variables of total sentences and words at an SNR of -3 dB in multi-talker babble (total administration time of 5 minutes). AUC = 0.86.

Table 9. The observed and the predicted classifications between controls and all mild MS using a logistic regression based on a 5-minute speech-in noise task by the cutoff of 0.68

	Predicted		
Observed	**Control**	**All mild MS**	**% Correct**
Control	17	3	85.0%.
All mild MS	5	20	80.0%
Overall % correct			82.22%

Note: Sensitivity = 20/(20+5)% = 80%. Specificity = 17/(17+3)% = 85%. False positive = 3/(3+17)% = 15%. False negative = 5/(5+20)% = 20%.

3.4 Discussion

Quantifying functional deficits in early MS remains a considerable challenge in the field (226); yet it is paramount for tailoring disease-modifying strategies and evaluating potential therapeutic candidates. This study evaluated the sensitivity of SiN tasks to measure early sensory and cognitive changes in pwMS. Three main results emerged from the current study: (1) SiN tasks, particularly those with a multi-talker babble background, sensitively detected speech discrimination deficits in early and mild pwMS (median EDSS = 0) with normal hearing; (2) there were mild/moderate correlations between SiN metrics and standardized neuropsychological assessments which indicate that pwMS with lower functional scores also had poorer speech discrimination in multi-talker babble; (3) a quick five-minute task with words and keywords presented in multi-talker babble at a single SNR was 82% accurate in classifying mild pwMS (median EDSS = 0) from healthy controls. Together, this indicates that SiN tasks measure MS-disturbances on a scale order of magnitude more sensitively (≤ 20% reduction in speech intelligibility in babble for pwMS compared to controls) than standard EDSS steps at the early and mild stages of the disease.

SiN deficits depended on noise type: speech-weighted noise measured zero to modest MS impairments in discrimination, whilst multi-talker babble elicited significant MS impairments at almost all SNRs (except for floor and ceiling SNRs) for all linguistic stimuli. Speech-weighted noise is an energetic masker that diminishes target audibility only through masking and blending of acoustic signals at the periphery (180). In contrast, multi-talker babble elicits confusion because of its similarity to speech and its saliency which will involuntarily capture attention, a perceptual interference known as informational masking (180). Our findings that SiN performance was disrupted only in multi-talker babble and not in speech-weighted noise shows that the SiN difficulties in people with mild MS are not due to linguistic difficulties but must be due to cognitive disruption. Although speech discrimination in babble was impaired in early and late mild pwMS groups, we did not identify any subjective difficulties in daily life by our early, mild pwMS (see **Table 1**) using the self-report Auditory Attention and Distress Questionnaire (AADQ). The AADQ was developed by our group, and has identified changes in everyday life in auditory tasks in high-performing Autism Spectrum Disorder people (141) and in advanced stages of MS (refer to Chapter Two). We propose that deficits in SiN tasks in babble reflect a cognitive deficit that has not yet impacted auditory performance and processing in daily life settings in pwMS. The absence of complaints might reflect redundant auditory processing (10), either intrinsic (multiple parallel auditory CNS representations (166)) or extrinsic (syntactic and semantic cues, or multimodal information through (say) lipreading) (183). Further, early pwMS may use neural compensatory mechanisms to reduce or mask functional deficits (67).

Both early and late mild MS groups were significantly impaired on all SiN tasks presented in multi-talker babble compared to healthy controls, but no differences were revealed between the MS groups. Both physically (EDSS) and neuropsychologically (cognitive tests used here), the two groups were functionally very similar, despite the significant difference in disease duration. There is contradictory evidence regarding disease duration on cognitive profiles (227) and on the relationship between cognitive impairment and level of physical disability (228). Regardless, the functional preservation in the late MS group as measured by standardized clinical measures, is supported by the fact that early and late discrimination abilities within our SiN tasks were also very similar. The only task that

differentiated early and late mild pwMS was the loudness discomfort levels (LDL); a sensory test typically used to evaluate hypersensitivity to sound (229). Overall, late mild pwMS reported significantly more discomfort than controls and early mild pwMS to multi-talker and speech stimuli presented at various intensities (dB), suggesting that despite similar discrimination performances and subjective experiences in daily life, late mild pwMS had less tolerance to louder stimuli than early pwMS. This could have implications for social avoidance and fatigue in sustained exposure to such acoustic environments.

As noted above, SiN deficits were specific to the informational multi-talker babble masker and not the energetic speech-weighted noise masker. The role of auditory attention in informational masking is established (201, 230). In the context of the 'cocktail party' phenomenon described by Cherry (1953), attention involves an interplay of bottom-up salience and top-down attention: a listener can ignore other speakers in favor of a target speaker, but when salient information arises, such as the listener's name, attention switches to the new speaker involuntarily (212). Switching, inhibiting, and sustaining goal-directed auditory streams are important for SiN discrimination in difficult listening environments that contain distracting linguistic and phoneme interference (230). Complex attentional processes such as selecting, dividing, and alternating attention are most often impaired in MS – and CIS - while simple attention span remains generally intact (59, 61). It is possible that the attentionally-demanding, salient acoustic characteristics of babble exacerbated MS attentional deficits, resulting in greater speech discrimination impairment in multi-talker babble, but not in speech-weighted noise. Inhibitory control, is a component of attention known to compromised in MS; poorer inhibitory control may increase susceptibility to distraction by the informational background masker of babble (196, 231). Alternatively, poorer attentional selectivity could impact encoding of the periodic temporal structures of the sounds.

One strategy for a listener to analyse a complex acoustic scene is to use 'glimpsing'(179) to extract information in "a time-frequency region which contains a reasonably undistorted 'view' of local signal properties" (232). Natural masking speech signals have fluctuating levels, where listeners can use momentary reductions in masker energy where the target signal is more audible, in order to form a representation of the target stream embedded in noise (140). If the encoding of fine temporal structure is compromised, listeners cannot parse complex acoustic scenes into coherent streams or objects that can be attended to (233). It is possible that MS listeners failed to take advantage of momentary 'dips' in babble. Such fluctuations are not present in the steady state speech-weighted noise, therefore, the temporal resolution skills required to use glimpsing to distinguish speech from multi-talker babble may partly explain why it is such a potent masker for early MS. Notably, MS listeners perform more poorly than controls in a words-in-noise paradigm when a wideband background noise has randomised silent periods, but not for continuous noise(86). In either instance, when a target signal is degraded due to reduced auditory temporal abilities or attentionally demanding characteristics of multi-talker babble, MS listeners must place greater demands on finite cognitive processes to reconcile perceptual ambiguity– a demand that is further exacerbated in smaller SNRs. Of relevance here is Working Memory, the limited-capacity temporary storage system for active maintenance of information in the face of ongoing processing and distractions (234). In the cognitive hearing sciences, the Ease of Language Understanding (ELU) model (198) emphasizes the subtle balancing act between bottom up and top down aspects of language

processing and how and when working memory is engaged to support the active maintenance of acoustic information in adverse conditions (198). Multifaceted working memory processes integrate lexical and phonological memory stores to 'fill in the gaps' of degraded sounds or mismatches between the perceptual speech input and phonological representations stored in long term memory (198). Working memory representations are vital for accessing semantic and syntactic relations amongst words and sentences to construct meaningful and coherent speech, and could also serve as templates that guide behaviour and bias perceptual activity (235). Working memory impairments are widely reported in early MS (189), however, working memory is not a unitary construct but a complex of several levels of processing broadly categorized as: maintenance and manipulation. Some studies have reported that pwMS have problems associated with maintenance in working memory, whilst others have concluded that the primary deficit is at the level of the central executive which controls and manipulates the contents of the working memory stores (60). Further, information processing speed, the speed and efficiency with which information is processed and integrated with other cognitive processes for formation of a behavioural response(236, 237), is prominently slower in early MS and has been proposed to be the underlying factor in cognitive domain deficits such as working memory and attention (60, 238). Thereby, slowed information processing speed may in part contribute to a greater cognitive load experienced by MS listeners in SiN conditions compared to healthy controls.

PwMS with cognitive deficits are likely to struggle with SiN tasks, particularly when temporal processing deficits further degrade the signal and thereby exacerbate demand on top-down processes. Interpreting SiN discrimination deficits as a reflection of cognitive MS impairments is supported by our finding of significant negative correlations between SNRs at 50% discrimination accuracy and standardized neuropsychological performance in the PASAT, SDMT and CVLT. The PASAT is a complex test of mental arithmetic, attention, working memory, information processing speed and places a heavy load on executive control processes (220). The SDMT is also a measure of processing speed, however it is regarded as a more superior test to the PASAT due to greater sensitivity, psychometric validity, patient acceptance and ease of administration. As a consequence, the SDMT is utilized in all recommended cognitive batteries for pwMS(239). The CVLT-II is a measure of episodic verbal learning and memory (222); a test which is particularly sensitive in early MS as verbal memory deficits have been reported in pwMS with a mean duration of 1.5 years (240). The digit span test (DST- WAIS-IV) did not correlate with any of our SiN tests. This may seem surprising given that the DST is referenced as a standardized measure of memory; however, the demands on short-term memory capacity in the digit span test are likely to be minimal. The digit span test requires participants to repeat a series of digits of increasing length that are orally presented at a one digit per second rate in silence. Normal forward digit spans (seven +/- two digits) have been reported in amnestic patients with Alzheimer's disease, Korsakoff's syndrome (241) and early phase MS (<4 years since diagnosis) (209). Backwards Digit span requires different processes or strategies as the task demands require mentally reversing the perceived sequence (242), but are also normal in early MS (209). None of the discrimination tasks in speech-weighted noise correlated with any neuropsychological task, consistent with the idea that higher phonological mismatches occur in noise that provides informational masking as opposed to purely energetic, which results in more effortful processing mechanisms based on working memory to make the speech comprehensible. However, we must point out that a limitation of our interpretations is that only pwMS were tested for the neuropsychological correlates of speech perception. The

cognitive functions required for speech discrimination in noise should ideally be examined and compared between groups.

Words, compared to sentences, elicited a greater degree of discrimination impairment in MS listeners. It may be that mild pwMS can exploit contextual and semantic cues present in sentences that are absent when words are presented in isolation. Reading comprehension and general linguistic competence has been shown to aid SiN intelligibility (224, 243), however the relationship has been reported in tests utilizing sentences- but not syllables (224). Words in isolation are likely to provide relatively more 'bottom-up' acoustic-phonetic cues, in contrast to sentences (244). The role of linguistic cues in more difficult and longer sentences could be investigated in future studies. Finally, we note the efficacy of our logistic regression model in classifying minimally impaired pwMS from healthy controls using predictors of keyword and word discrimination in babble at an SNR of -3 dB. The capacity for SiN tasks to measure subtle deficits specific to early MS may reflect their dependence upon dynamic heterarchical interactions between central auditory processing and cognition. Acoustic analysis of complex auditory scenes entails both exquisite local neural timing and the integrity of diffuse, higher-level networks, revealing subtle changes in MS that are not reflected in the EDSS. We do not suggest that our SiN task might replace more conventional neuropsychological measures in MS but suggest that it might be employed as a screening tool for changes in cognitive function, due to its ease of use and speed. Our SiN task only takes five minutes to administer, and it is cost effective and non-invasive, these features are advantageous in a clinical setting.

Preface – Chapter Four

Speech-in-noise (SiN) tasks sensitively capture sensory and cognitive impairment at various degrees of severity in multiple sclerosis (MS) (**Chapter Two)**, as well as early and mild stages of the disease prior to the development of overt symptomology (**Chapter Three)**. Multi-talker babble was a particularly potent masker in early MS; it's potency largely attributed to both its attentionally capturing salient features and temporal-resolving complexity. Expanding on the latter feature, the current chapter will investigate the working hypothesis that speech discrimination is a particularly vulnerable function in MS as it requires precise neuronal timing, which is disrupted by demyelinating lesions that slow conduction and cause neuronal dyssynchrony. Exquisite timing is important for capturing rapidly changing acoustic transitions not just within speech itself, but also in conditions where speech originates from different locations.

In the chapters up to this point, all speech and noise were presented with no differences in binaural timing or level cues and therefore would have modelled speakers and maskers coming from directly in front. Consequently, timing cues have only been presented monoaurally; in other words, both ears received identical acoustic input, thereby degraded speech could be resolved with one ear. This study will increase the temporal-resolving demands of SiN tasks by including an additional timing cue which requires interaction of complementary inputs of the two ears to discriminate speech. The inputs to the ears will differ within the millisecond range; referred to as an interaural timing difference (ITD). ITD processing requires the integration of sounds from both ears in the brainstem (as opposed to combining before sound reaches the ear (monoaural)) and results in the perception of sound originating along the listener's horizontal plane in space. It's important to add that while ITD processing is important for the perception of sound in space, interaural level differences (ILD)- a small decibel difference between the two ears- is another important acoustic cue present in spatially separated speech.

The added complexity in temporal processing during the analysis of a spatialised auditory scene may exacerbate SiN deficits observed in **Chapter Three** further, and thereby improve the sensitivity of SiN tasks to capture deficits in early, mild MS. We point out that the majority of participants in this study who completed all spatial tasks form a large subset of the participants in **Chapter Three**.

The study presented is written up as a manuscript intended for publication but not yet submitted.

Chapter Four: Discriminating Spatialised Speech in Complex Environments in Multiple Sclerosis

Pippa Iva[1], Russell Martin[1], Joanne Fielding[2], Meaghan Clough[2], Owen White[2], Gustavo Noffs[3], Branislava Godic[1], Anneke van der Walt[2], Ramesh Rajan[4]

[1]Department of Physiology, Monash University, Melbourne, Australia
[2]Department of Neurosciences, Central Clinical School, Alfred Hospital, Monash University, Melbourne, Australia
[3]Centre for Neuroscience of Speech, University of Melbourne, Melbourne, Australia
[4]Neuroscience Program Biomedicine Discovery Institute, School of Biomedical Sciences, Monash University, Melbourne, Australia

Abstract

Background:

Multiple sclerosis (MS) is a multi-component disease where inflammatory and neurodegenerative processes disrupt wide-ranging cerebral systems, including auditory networks. Although cochlear hearing loss is uncommon, people with MS (pwMS) frequently present with deficits in binaural hearing, which involves integration of sound inputs to both ears, for using acoustic spatial localization and disambiguating important signals from competing sounds. Spatial processing deficits have been described in pwMS using localization tasks of simple tones presented in silence but have yet to be evaluated in realistic listening situations, such as speech emanating from various spatial locations within a noisy environment.

Objectives:

To investigate how pwMS discriminate speech appearing to emanate from different spatial positions, in background competing conversation.

Methods:

Pre-recorded everyday Bamford-Kowal-Bench (BKB) sentences from a standard list and speech stimuli from the coordinate response measure (CRM) were presented via headphones with virtual acoustic techniques used to simulate as if they originated from 0º, 20º and 50º on the azimuth plane around the listener. BKB sentences were presented with eight-talker babble, whilst CRM target phrases were presented with a masker phrase that had similar syntactic structure and spoken by a speaker of the same sex. Controls (n=20) and age-matched pwMS with mild severity (Expanded Disability Status Scale (EDSS) score $\leq$ 1.5; n = 23) were required to repeat the target BKB sentence or select the target CRM coordinates. Mildly affected pwMS also completed the Paced Serial Addition Test (PASAT) and a basic three alternative forced-choice spatial task of detecting interaural time differences (ITD; a binaural spatial cue) in noise bursts. Additional participants with greater disease severity; moderate (EDSS 2.5 – 4.5; n = 16) and advanced disability (EDSS 5 – 7; n = 8) also completed the spatialised speech in babble noise task. All participants passed a standard hearing evaluation.

Results:

PwMS still benefited from the presence of spatial cues, as speech intelligibility improved when target speech was separated from the masker by 20º and 50º azimuth compared to 0º (co-localized), however, moderate and advanced MS groups did not achieve spatial release from masking to the same extent as controls. Specifically, a one-unit increase in spatial separation increased the odds of discriminating the correct sentence from babble for controls by 5%, but only 3% for pwMS with moderate and advanced disability. Mildly affected pwMS achieved similar spatial release from masking as controls when sentences were presented in noise but displayed less of a spatial advantage when the masker was a single competing talker of the same sex. PASAT scores were significantly associated with discrimination scores in colocalized and separated conditions.

Conclusions:

Knowing the spatial location of a sound is particularly critical in a complex noisy environment, as spatial cues help to group ambiguous sound elements into coherent streams. Although pwMS were able to use spatial cues, they did not receive the same spatial release from noise as controls. This is the first study to investigate how pwMS navigate their acoustic surroundings and communicate in noisy social environments.

4.1 Introduction

Multiple sclerosis (MS) is a multi-component disease where inflammatory and neurodegenerative processes disrupt wide-ranging cerebral systems (71), including auditory networks (10, 11, 245). MS affects the myelin coating of nerve fibres, and since the loci of demyelinating lesions can occur at multiple central nervous system (CNS) sites and develop or regress at different rates between people with MS (pwMS), nearly any auditory symptomatology pattern may result (9). The central auditory system appears comparatively more vulnerable than the peripheral one (246), such that pwMS typically present with normal pure-tone thresholds (81), but abnormal centralised processes such as dichotic listening (simultaneous presentation of stimuli to both ears) (145, 151, 153), speech understanding in noise (unpublished Iva et al., 2020) (12) and processing cues to auditory space (123-125).

There is considerable evidence on MS-related difficulties in processing spatial acoustic cues. Determining the location of a sound source is best with the use of binaural detection based on the fact that the two ears receive slightly different information generated by the interactions of sound waves with the external physical shape and location of the ears, head and body (118). Along the frontal azimuth plane (the horizontal dimension), two acoustic cues are dominant for localization: i) interaural time differences (ITDs) and ii) interaural level differences (ILDs) (119) in the onset or features of ongoing sounds. In practical terms, any source displaced from the sagittal plane results in the ear closest to it receiving the sound slightly quicker and louder than the other ear, i.e., in non-zero ITDs and ILDs. For sounds in the vertical dimension, spectral composition or frequency profile changes caused by the size of the head and pinnae are vital for elevation judgements and resolving back-front confusions (119). Binaural inputs converge at the superior olivary nuclei (SOC) in the brainstem, and the medial (MSO) and lateral (LSO) superior olives are responsible for the initial analysis of disparities in time (within a few milliseconds) and level (within a few decibels), respectively (247). A small number of studies have examined how pwMS judge location in the azimuth plane (123-127, 248). In lateralization tests, where listeners are required to indicate where they perceive dichotic stimuli in their head, investigators classified abnormal performances by pwMS as pre-dominantly (i) centre-orientated, i.e various non-zero ITD sounds were perceived to be located directly in front (0^{o} azimuth) or (ii) side-orientated; perception of sounds were biased towards left/right side positions ($\pm$ 90^{o} azimuth) (123) with centre-orientated performance principally associated with caudal pontine lesions and side-orientated performance with lesions rostral to the SOC (10). Another tool used to evaluate use of spatial cues in pwMS was the 'Just Noticeable Differences' (JNDs) task that determines the smallest ITD or ILD that a participant can detect reliably. Levine et al. (1993) reported that in pwMS, JNDs for ITDs were most affected for high-frequency sounds (76.3% abnormalities), whilst ILDs remained largely intact by the disease process (10.5% for low frequency and 7.9% for high frequency) (126). The selective detrimental effect of MS upon neural timing was largely attributed to demyelination slowing conduction velocities and causing dys-synchrony in neurons. Given that pure-tone audiometry and ILDs are tasks not dependent on conduction speed, this explains why performances by pwMS in these tasks remain relatively intact (10).

One limitation of these seminal studies is that evaluation of the acoustic spatial precision of pwMS was done with trains of clicks, noise bursts or tones presented in silence. None looked at ecologically valid conditions like those where spatial cues are used to distinguish between speakers in noisy environments. This is an important aspect of MS to investigate to better understand daily-life communication difficulties in pwMS, since normal human speech communication routinely occurs in social, learning and work environments that are noisy. Especially in such situations, spatial cues help a listener to tease apart ambiguous sound elements into coherent auditory objects and selectively attend to sounds arriving from one direction whilst ignoring sounds arriving from another (249). The 'cocktail-party' phenomenon (212) is a common scenario that confronts many listeners with competing dynamic acoustic stimuli (speech or non-linguistic sounds) originating from various locations in three-dimensional space (180, 212). Listeners use a combination of strategies to increase the intelligibility of a target speaker under unfavourable signal-to-noise ratios (SNR: the ratio of levels of target and interfering sounds); one of the most advantageous acoustic methods is perceiving the spatial configuration of sound sources (249, 250). Several studies have demonstrated that speech recognition in noise improves when the source of the speech is separated horizontally from the interference (249, 251-255). This improvement in speech intelligibility is referred to as a spatial release from masking (SRM) (252), and can be calculated as the difference in scores of two test conditions of a speech-in-noise (SiN) task, where the only difference between the two conditions is the spatial separation between the origin of the target and the origin of noise (256). Acoustic cues which affect the magnitude of SRM are largely head shadow and binaural interactions. If the masker is closer to one ear, the head's acoustic 'shadow' creates a masker ILD which will increase the SNR at the other ear (118, 257); a listener can choose to move their head or body to orient themselves to create and take advantage of such a 'better ear effect'. The other advantage depends on the ability of the auditory system to utilize both ITD and ILD cues from the signal, also known as the binaural squelch effect (118). Bronkhorst and Plomp (1988) concluded that the effects for ILDs were larger at high frequencies, and ITDs had the greatest effects at lower frequencies (257). Effective use of spatial processing can increase speech intelligibility by as much as 12 dB in adults depending on the condition (258), (i.e participants understood the target speech at the separated condition at a 12 dB reduction in SNR compared to colocalization). It is worth noting that a distinction can be made for ITD and ILDs cues present in sounds located within the peripersonal space (approximately one metre from the listener) and farther sounds in extrapersonal space, i.e. ILDs for low frequency sounds are large for sound sources within the peripersonal space but very small for sounds beyond that (259). Given that MS lesions have a predilection for brainstem sites (260), the location where binaural interaction initially occurs to interpret ITD and ILDs (261), pwMS may potentially have deficits in spatial processing ability and reduced SRM. This hypothesis remains to be investigated.

The aim of our study was to evaluate how pwMS discriminate speech appearing to emanate from different spatial positions along the frontal azimuth plane, in a background of competing conversation. This is the first study to investigate how pwMS use spatial cues to navigate their acoustic surroundings and communicate in noisy social environments. It is rarely the case that the source of a target speech signal and the source of a competing masker are perfectly co-located in space, therefore, this study reflects an important aspect of successful speech processing in everyday life.

4.2 Methods

All procedures were approved by the Monash University Human Research Ethics Committee (8170) and Melbourne Health HREC (2015.069). The study conformed to guidelines of the National Health and Medical Research Council of Australia and the Helsinki Declaration protocols for experiments involving human participants.

4.2.1 Participants

This cross-sectional study evaluated the auditory performance of 20 controls and 47 pwMS confirmed by revised McDonald criteria (172). The pwMS were recruited through Royal Melbourne Hospital Australia, and neurologically healthy controls were recruited from the local community. All participants provided informed written consent. PwMS were grouped according to the Expanded Disability Status Scale (EDSS) score (171) as rated by a neurostatus certified neurologist at study entry, which was within six months of the auditory testing reported here. PwMS with EDSS scores ≤ 1.5 were classified as 'mild'; between 2 – 4.5 as 'moderate' and between 5 – 7 as 'advanced' disability.

Exclusion criteria for both MS and control participants were a history of another neurological disorder, substance abuse/dependence, pregnancy, and/or the presence of hearing loss (see section 2.3 Audiometry). No pwMS experienced exacerbated symptomology for at least three months prior to this study. All participants reported English as their native language, which has been shown to be important for SiN tasks in previous work (262-264).

4.2.2 Neuropsychological testing

Only mildly affected pwMS participated in the Paced Auditory Serial Addition Test (PASAT) (220) as part of a routine checkup with a neurologist. The PASAT tests were done within a year of the auditory testing reported here.

4.2.3 Audiometry

The hearing status of all participants was determined using pure tone audiometry with a Beltone Model 110 Clinical Audiometer and calibrated TDH headphones. Hearing sensitivity was tested, first in one ear and then in the other, at the following frequencies: 250Hz, 500Hz, 750Hz, 1000Hz, 1500Hz, 2000Hz, 4000Hz, 6000Hz and 8000Hz. A pure tone was presented to the test ear through headphones and participants were required to press a hand-held button to indicate the sound was heard. Hearing thresholds were recorded in decibels Hearing Level (dB HL) relative to normal hearing sensitivity (ISO 8253-1, 1989) were defined as the lowest intensity in decibels (dB) at which the tone is perceived 50% of the time. As is standard in audiology, a modified Hughson-Westlake procedure (174) was used to determine hearing thresholds. Pure tone averages of the hearing thresholds levels at 500, 1000, 2000 and 4000 Hz were obtained for all participants.

4.2.4 Interaural time differences

A computer-controlled, three-alternative, forced-choice paradigm was used to determine the ability of listeners to detect ITDs in broad band noise bursts presented binaurally. Two of the three identical noise bursts in each presentation

had no ITD between the left and right ear, ideally resulting in the listener perceiving these noise bursts to be located directly in front (0º azimuth). One of the three noise bursts contained a small ITD between the left and right ears. This noise burst to contain the ITD was randomly allocated to a different ordinal position of the three noise binaural noise bursts in each presentation and on the basis of the ITD, the participant had to state which of the three binaural noise bursts they perceived to be slightly different to the other two (i.e. which noise burst is presented with an ITD). The participant indicated their choice by clicking on one of three numbered buttons corresponding to the three noise bursts presented on a computer screen.

Stimuli were presented through Sennheiser HD535 headphones at 80 dBA (A-weighted decibels). Based on parameters previously used in previous studies of ITD discrimination in pwMS (122, 126, 127), the noise bursts (white noise filter band passed between 0.2 - 1 kHz) were 275 ms long with 800 ms inter-stimulus intervals. Six different ITDs (1000, 500, 250, 100, 50, 10 μs) were tested, each being presented 20 times to the listener in a randomized-block design to result in a total of 120 trials. Participants were given a trial run for noise bursts with an easily-detectable ITD of 3 ms so that they were familiar with the task.

4.2.5 Spatial speech-in-noise task

The stimuli used for the spatial speech-in-noise (SSiN) task have been detailed previously (141, 175). In brief, speech stimuli were derived from a standard clinically-used battery of sentences, the Bamford-Kowal-Bench (BKB) sentence lists (176) consisting of simple sentences in common use (176), each containing 4-6 words with 3 being keywords for scoring. Sentences were presented simultaneously with multi-talker babble, which consisted of eight simultaneous voices generated by doubling over and temporally offsetting a recording of four people reading nonsense text.

Speech and babble were presented binaurally through Sennheiser HD535 headphones. Thirty BKB sentences were randomly allocated to three blocks that were equalized, based on previous laboratory data (265), for detectability in noise conditions with no spatial modifications. Each block was then filtered through the human-head related transfer function (HRTF) to create sets that appeared to be emanating from the headphones at 0º, 20º and 50º in azimuth on the listener's interaural horizontal plane (**Figure 1A**). A generic HRTF was generated (266) to modify the sentences via direct measurement of direction-related information at the ears with two microphones. Each location consisted of 10 unique target sentences that the listener had to verbally repeat back to the experimenter or indicate that they were unable to do so. Correct responses were defined as all three keywords being correctly identified in the correct order. Multi-talker babble was presented at 70 dBA at equal levels in both ears, thus appearing to emanate from directly in front of the listener. Stimuli were presented at a fixed SNR of -3 dB as it elicited ~50% discrimination performance at no spatial separation (pre-determined based on a pilot study).

No time limit was placed for a response and feedback was not provided back to the listener. The experimenter recorded the appropriate response that was repeated by the listener in an in-house program and another randomized sentence was presented after a 1.5 second delay.

4.2.6 Coordinate response measure task

The coordinate response measure task (CRM) is a non-standardized communication performance task adapted from similar tasks by Moore (1981) as a measure of speech intelligibility. The general procedures for the CRM task are similar to previous studies (267).

Two sentences were presented simultaneously to the listener in a randomized-blocks design over 90 trials. These sentences were from a publicly available, modified version of the Air Force Research Laboratory's CRM speech corpus described by Bolia et al. (2000). The sentences in the CRM consist of a call sign and a colour number combination embedded within a carrier phrase. Each sentence has the same rigid syntactic structure: "Ready [callsign] go to [color] [number] now," where the sentence can be any factorial combination of eight call signs ("baron", "arrow", "charlie", "eagle", "hopper", "laker", "ringo", "tiger"), four colours ("blue", "green", "white", "red"), and numbers (1 to 8). Combinations yield a total of 256 phrases, all of which were recorded by eight talkers (four males and four females between the ages of 18 and 26), for a total of 2048 unique phrases. The sentence recordings in the publicly available CRM corpus were band-pass filtered from 200 Hz to 18 kHz and trimmed to remove extraneous clicks and silence before and after the speech signals. The average message duration was 1.8 seconds and presented via Sennheiser: HD 535 headphones at 70 dBA.

The target sentence always addressed the callsign "Baron," but the color and number it referred to and its talker varied randomly from trial to trial. The masker sentence is chosen at random with the constraint that it addressed a callsign other than "Baron," refer to a color and number different from those referred to in the target sentence, and be spoken by a talker different from, but of the same sex as, the target sentence's talker.

Within each session, all possible combinations of target location (i.e., -50°, -20°, 0°, +20°, +50° azimuth) and masker location (i.e., -50°, -20°, 0°, +20°, +50° azimuth) to create three spatially separated conditions (0°, 20° and 50°) (**Figure 1B**) were sampled an equal number of times in random order. As a result, the participant was uncertain as to which location the target would be presented from on any given trial. The participant's task was to indicate the colour and number referred to in the target sentence by clicking a cursor on a computer monitor to the appropriate co-ordinate in a matrix of coloured numbers (see **Figure 1B**). A response was regarded as correct if both the colour and the number referred to in the target sentence were correctly identified. Participants were given no time restrictions and clicked on a "next" button whenever they felt prepared to hear the next trial.

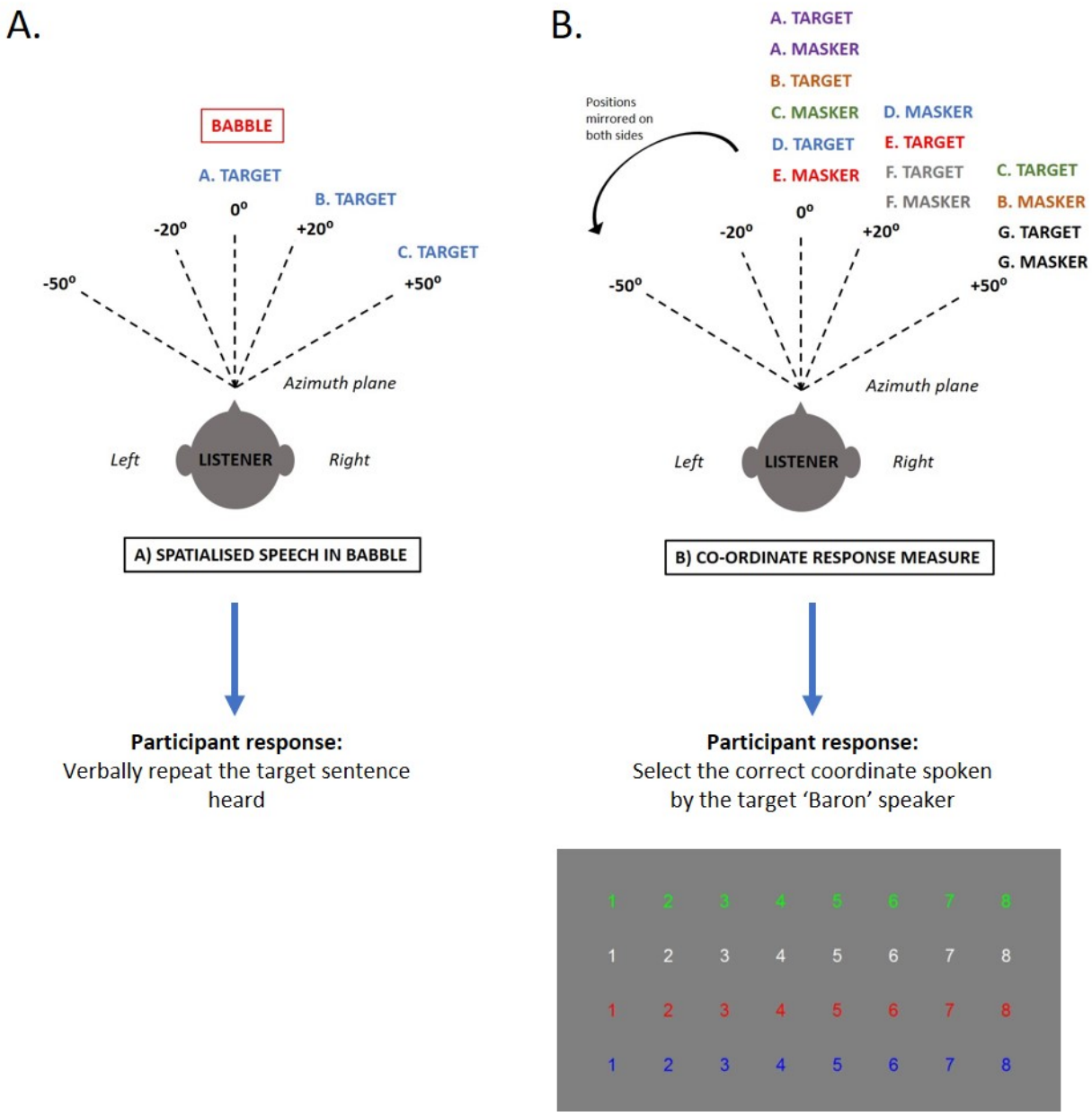


Figure 1. A schematic of the spatialised conditions for the speech in babble noise (SSiN) (A) and coordinate response measure (CRM) task (B). In the SSiN (A), babble was always perceived directly ahead (0º azimuth), whilst ten different Bamford-Kowl-Bench (BKB) target sentences were perceived at each of the three locations: 0º, +20º or +50º azimuth. Participants were required to verbally repeat the sentence heard. In the CRM (B), target and marker locations could be located at -50º, -20º, 0º, +20º or +50º azimuth to create three spatially separated conditions (0º, 20º and 50º), (13 different configurations, although schematic only shows seven (A-G); the other six are mirrored on the other side. Participants were required to listen for the coordinates spoken by the target speaker and click the corresponding coloured number on a computer screen

4.2.7 Statistical analyses

Statistical analyses were run using IBM SPSS Statistics 26, MATLAB 2019b and GraphPad Prism 8 programs.

Participant demographics and hearing sensitivity (pure-tone averages) were compared between control, mild, moderate, and advanced MS groups by Chi-squared tests, Kruskal-Wallis Tests and One-Way ANOVAs, depending on the distribution of data sets.

All spatialised tasks (i.e. ITD discrimination, SSiN and CRM) were evaluated using two-way mixed-effects analysis of variance (ANOVA) and appropriate post hoc Tukey's multiple comparisons tests.

To identify factors that significantly influenced speech discrimination accuracy on any given trial (0=incorrect; 1=correct), binomial generalized linear mixed effects models (glme) with logit link functions were generated for each spatially separated condition of the CRM task to determine how fixed effects of age (years), disease duration (years), ITD discrimination (average % correct) and PASAT score (% correct) influenced release from masking (0=incorrect; 1=correct on any given trial) in mildly affected pwMS.

Potential fixed-effects were explored with a participant-specific random intercept representing between-participant heterogeneity. All variables had variance inflation factors (VIF) <1.35, well below the recommended cut off VIF of 5, indicating no problematic levels of multicollinearity among predictors.

4.3 Results

4.3.1 Participant demographics, characteristics, and audiometric hearing status

20 controls and 23 mildly affected pwMS (EDSS ≤ 1.5) were evaluated in this cross-sectional study. For participant details, refer to **Table 1**. Participants evaluated in the current study had bilaterally normal hearing between 250 – 4000 Hz; except for 5% of participants from each group who had small hearing losses (of 5–10dB) at the higher frequencies of 6000 and 8000 Hz in one ear only. Potential confounders such as age, and the four-frequency pure tone average (the average hearing threshold across 500, 1000, 2000 and 4000Hz in each person; a standard audiometric measured to index a person's hearing) (dB HL) in the left and right ears were not statistically different between controls and the mild MS group ($p>0.05$).

Table 1. Participant characteristics

		Control	Mild MS	*p*
Demographics	**Number of participants**	20	23	
	Sex F(M)	17(3)	24(3)	
	Age, (yrs)			
	Mean (SD)	45.85(10.82)	47.03(8.73)	0.99 [a]
	Range	28 - 60	28 - 65	
Audiometry	**Pure tone average (dB HL)**			
	Left (Mean, SD)	13.12(5.90)	12.46(4.67)	0.68 [b]
	Range	-1.25 to 18.75	2.5 to 18.8	
	Right (Mean, SD)	12.06(5.13)	13.59(5.37)	0.35 [b]
	Range	2.5 to 23.75	2.5 to 22.5	
Disease Characteristics	**Disease duration (yrs)**	-		
	Mean (SD)	-	10.45(5.69)	
	Range	-	1 - 22	
	***EDSS**	-		
	Mean(SD)	-	0.11(0.37)	
	Range	-	0 - 1.5	
	Phenotype RR(SP)		23(0)	
	On disease modifying therapy (n, %)	NA	91.3%	

F = female; M = male
SD = Standard deviation
dB HL = decibels hearing level
*EDSS = Expanded Disability Status Scale Score
RR = Relapsing Remitting
SP = Secondary Progressive

[a] Mann-Whitney Test
[b] Unpaired t-test

4.3.2 Discrimination of interaural time differences

An investigation was made into how MS listeners discriminated ITDs embedded in noise bursts presented in silence. This task did not incorporate complex speech materials, or the type of auditory scene analysis required to parse competing sound sources as its purpose was to establish how pwMS simply detected ITDs. Our findings suggest that mild MS did not affect the ability to detect ITDs in noise bursts. Mean percentage (± standard error of the mean (SEM)) of correctly discriminated ITDs (10-1000 μs) by control (n = 20) and mildly affected pwMS (n = 20; missing data from three participants) are presented in **Figure 2**. Both controls and mildly affected pwMS detected ITDs greater than 250 μs quite well (≥ 95%), however, discrimination performance declined for all listeners as ITDs became shorter and hence, more difficult to detect. Chance level of ITD discrimination is represented in the graph as a dotted line (33.33%, i.e 1 in 3). Visually, the graph shows similar discrimination performances by pwMS and controls for most ITDs, but with a slight divergence in discrimination ability at the smallest time difference of 10 μs. These effects were tested using a 2 x 6 [(control and mild MS)] x (ITD = 1000, 500, 250, 100, 50, and 10 μs)] two-way mixed ANOVA. As expected, ITD influenced discrimination ability for low band pass noise [$F(5,190) = 188.3$, $p < 0.0001$, $\eta^2 = 71.93$]. There was no listener group effect [$F(1,38) = 0.36$, $p = .55$, $\eta^2 = 0.13$] or interaction effect [$F(5,190) = 0.41$, $p = 0.84$, $\eta^2 = 0.16$]; suggesting that mildly affected pwMS in this group have similar ITD detection to controls, overall.

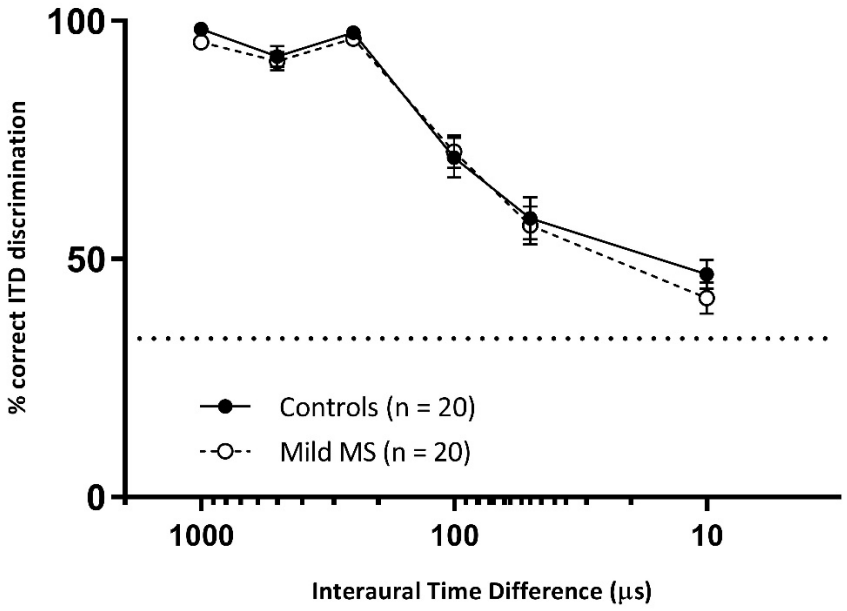


Figure 2. Mildly affected pwMS have similar interaural time difference (ITD) detection thresholds to age-matched controls. Mean percentage (± SEM) of correctly discriminated ITDs (10-1000 μs; logarithmic scale) by controls (n=20; black circle) and MS listeners (n=20; open circle). The dotted line represents chance level (33.33%) for the three-alternative, forced-choice paradigm. Performances by controls and MS listeners were not significantly different on a group level. Two-way mixed ANOVA (p>0.05).

4.3.3 Spatial release of speech from multi-talker babble noise

Participants completed a spatialised version of our SiN task (265). In this task, controls (n = 20) and mildly affected pwMS (n = 23), were presented with sentences binaurally with the sentences being perceptually located at 0º, +20º and +50º azimuth through the incorporation of virtual auditory spatial cues (266). At the same time multi-talker babble noise was simultaneously presented at a perceptual location of 0º in azimuth. At each location, the SNR was fixed to be -3 dB, and we determined how many sentences the participants were able to detect correctly out of ten sentences. All three test conditions were interleaved randomly so order effects were eliminated.

Performance (mean sentences ± SEM correctly discriminated) by controls and mildly affected pwMS at each spatial separation are presented in **Figure 3**. Spatial separation of speech from noise improved sentence intelligibility for both listener types, as evident by increases in the number of sentences correctly identified with increasing spatial separation from 0º to 20º and 50º separation. Specifically, controls improved by ~ 38%, whilst pwMS (mild) improved by ~ 34% when sentences were separated from babble by 50º in azimuth compared to when speech was co-localised.

A 2 x 3 [(control, mild MS)] x (0º, 20º, 50º separation)] two-way mixed ANOVA confirmed that there was no significant interaction effect [$F (6, 82) = 0.74$, $p = .48$, $\eta^2 = 0.39$], but there was a significant effect of spatial separation [$F (2, 82) = 93.99$, $p < .0001$, $\eta^2 = 50.12$] as well as listener group [$F (1, 41) = 10.17$, $p < .01$, $\eta^2 = 22.36$] on sentence discrimination. Together, these results indicate that pwMS (mild) achieved similar spatial release from masking, but discriminated fewer sentences correctly compared to controls.

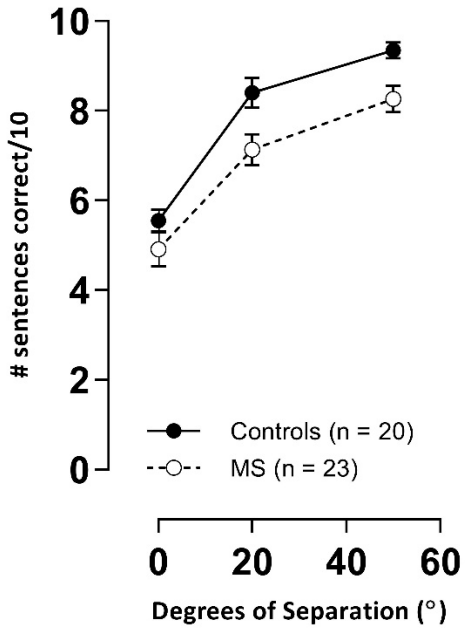


Figure 3. Mildly affected pwMS discriminated significantly fewer sentences in babble overall but achieved similar spatial release from masking compared to controls. Mean (± SEM) sentences correctly discriminated at 0º, 20º and 50º separation in azimuth from multi-talker babble by mildly affected pwMS (EDSS 0-1.5; n = 23; open circle/broken line) and controls (n = 20; closed circle, solid line). Two-way mixed ANOVA (p>0.05)

4.3.4 Spatial release of target speech from masker speech in the coordinate response measure task

Utilising different speech materials but the same spatial separations as the SSiN task, participants also completed the CRM task in which participants had to follow instructions from only one of two simultaneous speakers separated by three virtual azimuthal spatial separations (0º, 20º and 50º azimuth). The two speakers stated different colour-number coordinates for a colour-number grid on a PC screen and the participant had to pick the coordinate tagged by a cue in the speech of one speaker randomly. Mean ± SEM (%) target coordinates correctly identified are presented in **Figure 4**. SRM was observed in that response accuracy (the % of correct coordinates) improved in the spatially separated conditions (20º and 50º) compared to co-localised speech (0º), and this was found in both controls and pwMS. Controls improved by ~28% when speech was separated by our maximum virtual separation of 50º on the azimuth plane compared to when speech was co-localised, whilst pwMS improved by ~17.8%.

A 2 x 3 [(control, mild MS)] x (0º, 20º, 50º separation)] two-way mixed ANOVA confirmed the effect of spatially separated speech and listener group on discriminating CRM coordinates. There was a statistically significant interaction between the effects of spatial separation (º) and listener group on coordinate discrimination [$F(2,82) = 6.12$, $p = .004$, $\eta_p^2 = 1.69$], but no main effect of treatment [$F(1,41) = 0.95$, $p = 0.34$, $\eta_p^2 = 1.15$].

Figure 4 visually shows that the difference in mean ± SEM (% correct discrimination) between control and mild MS groups increased with greater spatial separation. At 50º spatial separation, pwMS discriminated 80% of coordinates accurately whilst controls discriminated 89%. A Sidak's multiple comparison test confirmed that pwMS discriminated significantly fewer coordinates than controls at 50º ($p = 0.04$).

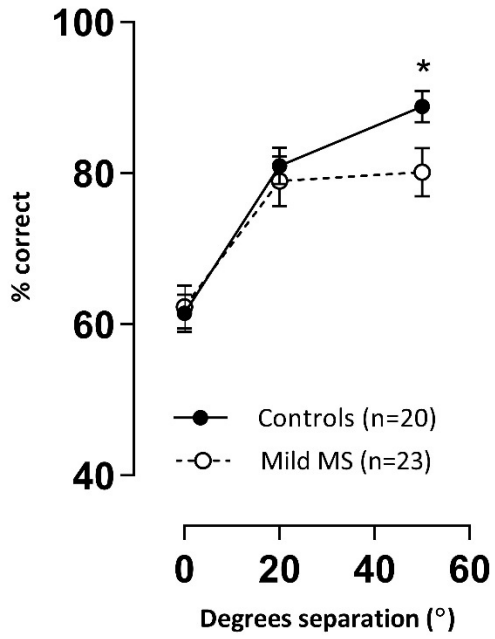


Figure 4. Mildly affected pwMS had less spatial release from masking in the coordinate response measure task compared to controls. Mean (± SEM) target coordinates correctly discriminated from masker coordinates at spatial separations of (0º, 20º and 50º) for pwMS with an EDSS ≤ 1.5 (n = 23; open circle, broken line) and controls (n = 20; closed circle , solid line). * $p < 0.01$; two-way mixed ANOVA with Sidak's multiple comparisons test

4.3.5 Multivariate analysis on factors influencing spatial release from masking in the coordinate response measure task

To explain the impact of multiple factors influencing spatial release from masking in the CRM task, we built a model to incorporate variables theoretically important for explanatory power. Spatial separation in azimuth (º), disease status (healthy vs. mild MS), age (yrs), hearing sensitivity (dB HL), and trial order effects (1 to 90) were fixed effects included in the regression model. To test our main hypothesis that pwMS do not achieve the same degree of spatial release from masking, an interaction between disease status and spatial separation was included as a fixed effect. Interaction terms (viz., spatial separation*age and spatial separation*hearing sensitivity) were also theoretically important to include, as their influence on spatial release from masking effects has been previously demonstrated(268).

A generalized mixed-effects model (glme) was generated using MATLAB Statistic Toolbox Release 2019b. Parameter estimates of fixed effects are listed in **Table 2**, along with the t statistic, degrees of freedom and p-values for each fixed effect to test the null hypothesis that the coefficient equals zero. The odds ratio (OR) and confidence intervals (CI,95%) are also included to quantify the magnitude of the association between the fixed effect and the outcome.

Beta coefficients for hearing sensitivity, age and their respective interactions with spatial separation were not significantly different from 0 ($p < 0.05$), thereby suggesting that they did not influence coordinate discrimination (0 = incorrect; 1 = correct). Fixed effects of spatial separation, trial order and an interaction term between spatial separation and disease status were significantly associated with coordinate discrimination in the CRM task ($p<0.05$). To quantify their associations, the ORs were interpreted. First and foremost, when all other factors are constant, a one-unit increase in spatial separation increased the odds of discriminating the correct coordinate for controls by 5%, but only 4% for mildly affected pwMS (1% lower compared to their healthy counterparts; $(1- e^{-0.01})$). Thereby confirming that pwMS do not receive the same spatial advantage as controls. There was also a positive association of trial order on CRM performance as the odds of correctly identifying the target coordinate was 1.02 (95% CI, 1.01-1.02) times greater compared to the previous trial.

Table 2. Parameter estimates of fixed effects of the generalized linear mixed effects model (with a logit link function) used to predict correct coordinate discrimination on each trial

						95% C.I for e^{β}	
	Estimate				OR		
Name	**(β)**	**SE β**	**tStat**	***P***	**(e^{β})**	**Upper**	**Upper**
Intercept + $N(0, \sigma_b^2)$	0.22	0.54	0.40	0.69	1.24	0.43	3.60
Trial order	0.02	0.00	11.01	**<0.0001**	1.02	1.01	1.02
Spatial separation (º)	0.05	0.01	4.83	**<0.0001**	1.05	1.03	1.08
Hearing sensitivity (dB HL)	-0.01	0.02	-0.37	0.71	0.99	0.95	1.04
Age (years)	-0.01	0.01	-0.40	0.69	0.99	0.97	1.02
Disease Status (ref: controls)							
Mild MS (EDSS ≤ 1.5)	0.12	0.22	0.55	0.59	1.13	0.74	1.72
Interaction terms (with spatial separation (º))							
Spatial separation x Hearing	0.00	0.00	-1.31	0.19	1.00	1.00	1.00
Spatial Separation x Age	0.00	0.00	-0.87	0.38	1.00	1.00	1.00
Spatial Separation x Mild MS	-0.01	0.00	-3.31	**<0.001**	0.99	0.98	0.99

Outcome: 0 = incorrect and 1 = correct coordinate
Degrees of freedom = 3861; Based on 3870 observations (N = 43 participants);
The estimate of the variability of the random-intercept effect (σ_b^2) = 0.60
Ordinary R^2 = 0.15; Adjusted R^2 = 0.14
SE = standard error
OR = Odds ratio
C.I = confidence interval
dB HL = Decibels hearing level
EDSS = Expanded Disability Status Scale

4.3.6 Multivariate analysis on disease factors influencing performance at each spatial condition in the coordinate response measure task

To examine the significant interaction effect in the CRM task (as confirmed by a two-way repeated measures ANOVA), a separate analysis was conducted on the mild MS group (the group tested in the CRM) at separate spatial conditions (0º, 20º and 50º separation). Three generalized mixed-effects models (glmes) with logit link functions were built to incorporate theoretically important fixed-effects variables: age (years), disease duration (years), hearing sensitivity (db HL), ITD discrimination at the lowest threshold tested (10μs) (%), PASAT scores (%) and trial order effects (1 to 30) to determine what influenced CRM performance. **Table 3** illustrates the amount of variance accounted for, coefficient estimates of fixed effects and their p values. Although trial order and intercept values were significant predictors in all models, they were not presented in **Table 3** as they were not variables of interest.

PASAT performance was a significant predictor at all three spatial conditions whilst ITD discrimination performance was only significant at the largest spatial separation of 50º ($\beta = 0.028$, $p < 0.01$). The coefficients for age and disease duration did not significantly contribute to the models for any spatial condition.

Table 3. Generalised Linear Regression Models predicting spatial release from masking at different spatial separations

Spatial Separation	R^2	Age (years)		Disease Duration (years)		ITD discrimination ability (%) at 10µs		Paced Auditory Serial Addition Test	
		Coefficient estimate	P value	Coefficient estimate	P value	Coefficient estimate	P value	Coefficient estimate	P value
0	9.2	-0.018	0.225	-0.010	0.511	0.009	0.250	**0.019**	**<0.0001**
20	27.1	-0.039	0.158	-0.016	0.692	0.016	0.251	**0.031**	**<0.001**
50	12.6	-0.019	0.340	-0.010	0.383	**0.028**	**0.009**	**0.045**	**<0.0001**

Outcome: 0 = incorrect and 1 = correct coordinate
Each model had 594 degrees of freedom and were based on 600 observations (N = 20 participants; 30 observations each);
The estimate of the variability of the random-intercept effect ($\boldsymbol{\sigma_b^2}$): Model 1 (0º) = 0.17; Model 2 (20º) = 0.52; and Model 3 (20º) = 0.37
Model 1 (0º) Ordinary R^2 = 0.09; Adjusted R^2 = 0.08;
Model 2 (20º) Ordinary R^2 = 0.27; Adjusted R^2 = 0.26
Model 3 (50º) Ordinary R^2 = 0.12; Adjusted R^2 = 0.11
ITD = interaural time difference

4.3.7 Errors in the coordinate response measure task

An investigation into the types of errors made by pwMS (EDSS ≤ 1.5) and controls in the CRM task was conducted. Error analysis can provide additional insights as to how listeners complete the task, and has been previously done in studies using the CRM speech corpus (269, 270). An error can be made by answering the number and colour from the masker sentence (MM), a combination of both masker and target, or an answer that neither comes from the target nor masker (//). Specifically, there are eight types of errors that can be made for each trial: 1) MM = represents the colour and number from the masker; 2) MT = colour from masker and number from target; 3) T/ = colour from target and number from neither target nor masker; 4) TM = colour from target and number from masker; 5) // = both colour and number are from neither masker nor target; 6) /M = colour from neither masker nor target and number from masker; 7) /T = colour from neither masker nor target and number from target; 8) M/ = colour from masker and number from neither masker nor target.

The mean ± SEM (%) of error types were calculated for each participant and averaged across each of the pwMS and control groups (**Figure 5)**. The most common error was responding with the coordinates spoken by the masker voice (MM; 44.99 ± 1.7% for pwMS and 43.65 ± 1.6 % for controls). This demonstrated that all listeners were able to separate target elements from each other (segregate) but made errors when attributing the coordinates to the correct stream. MT was the next most common error, followed by TM. All the other combinations made up ≤ 10% of the total errors combined. A chi-squared test of independence showed that there was no significant association between disease status and error type, X^2 (7, 43 = 6.75, p = 0.46); pwMS and controls made similar error types throughout the CRM task. Thus, while mildly affected pwMS did not get the same spatial release from masking for the largest spatial separations between the cued voice and the competing voice, their performance errors were similar in number and type to those made by controls. This suggests that the basic processes of making the decision as to which voice to respond to was the same between people with mild MS and controls.

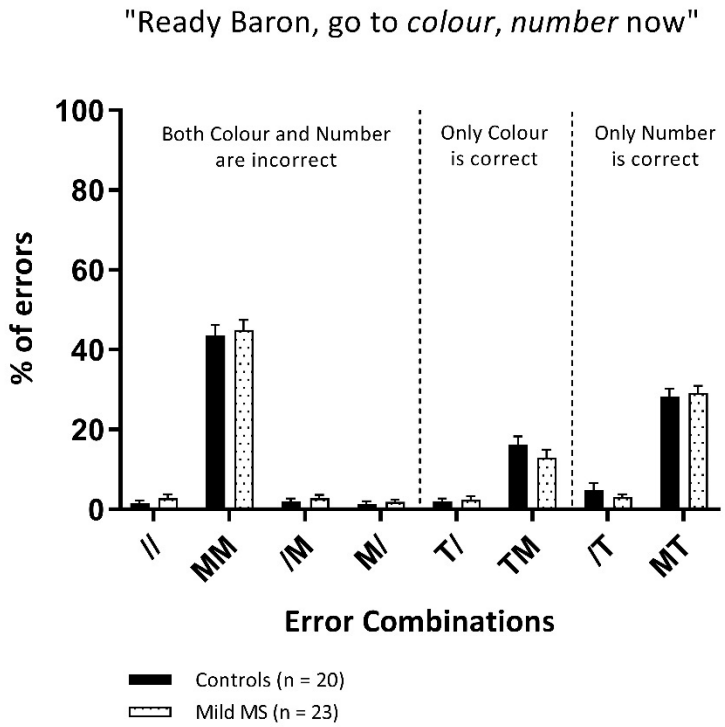


Figure 5. Mildly affected pwMS and controls made similar error combinations in the coordinate response measure task. Mean percentage (±SEM) of errors made by controls (black) (n = 20) and mildly affected pwMS (patterned) (n = 23) in the CRM task. // = both colour and number are from neither masker nor target; MM = colour and number both from the masker; /M = colour from neither masker nor target and number from masker; M/ = colour from masker and number from neither masker nor target; T/ = colour from target and number from neither target nor masker; TM = colour from target and number from masker; /T = colour from neither masker nor target and number from target; MT = colour from masker and number from target. (Chi-squared test, $p > 0.05$)

4.3.8 Impact of disease severity on discriminating spatialised speech in babble

For the final part of the results section, we will present additional data from the SSiN task described in **Section 3.3** obtained from participants with a wider spectrum of severity (EDSS scores between 2 – 7). PwMS with EDSS scores between 2 – 4.5 were classified as 'moderate' and between 5 – 7 as 'advanced' disability. In total, 16 moderate (EDSS 2- 4.5) and 8 advanced (EDSS 5 - 7) MS participants in addition to the 20 controls and 23 mildly affected pwMS previously described completed the task. For participant details, refer to **Supplementary Table A.4.**

Performance (mean sentences ± SEM correctly discriminated) by controls, mild, moderate and advanced MS listeners at each spatial separation are presented in **Figure 6**. Once again, spatial separation of speech from noise improved sentence intelligibility for all listener types, as evident by increases in the number of sentences correctly identified with increasing spatial separation from 0º to 20º and 50º separation. Specifically, controls and pwMS with mild, moderate and advanced disability improved by ~ 38%, 34%, 23% and 31% respectively when sentences were separated from babble by 50º in azimuth compared to when speech was co-localised.

A 4 x 3 [(control, mild, moderate and advanced MS)] x (0º, 20º, 50º separation)] two-way mixed ANOVA confirmed that there was no significant interaction effect [$F (6, 126) = 0.92$, $p = .48$, $\eta^2 = 1.14$], but there was a significant effect of spatial separation [$F (2, 126) = 66.84$, $p < .0001$, $\eta^2 = 27.60$] as well as listener group [$F (3, 63) = 15.65$, $p < .001$, $\eta^2 = 16.72$] on sentence discrimination. Simple main effects analysis confirmed that mild, moderate and advanced MS patients discriminated fewer sentences than controls ($p = 0.03$; < 0.01 and <0.0001, respectively). There was no difference in discrimination between mild and moderate MS groups ($p = 0.86$). The advanced MS group, however, did significantly worse than both the mild ($p < 0.0001$) and moderate groups ($p < 0.01$).

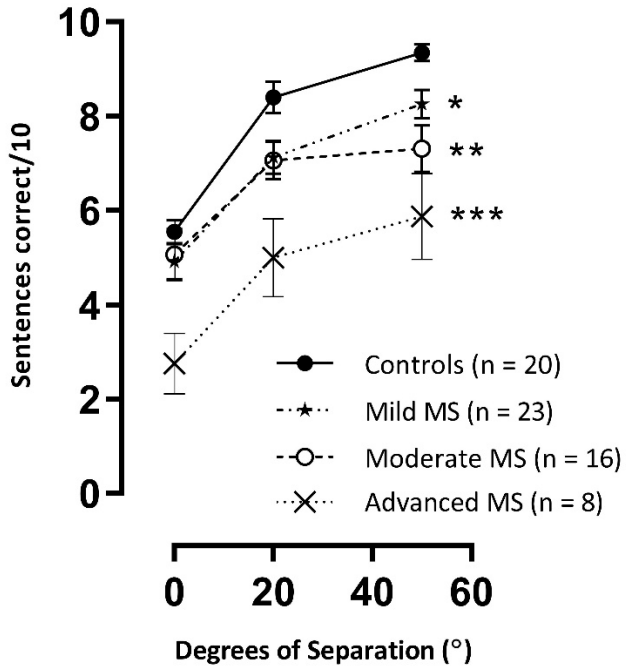


Figure 6. MS participants with mild, moderate, and advanced disability discriminated significantly fewer sentences in babble overall, compared to controls. Mean (± SEM) sentences correctly discriminated at 0º, 20º and 50º separation in azimuth from multi-talker babble by mild (EDSS 0-1.5; n = 23; star/broken line), moderate (EDSS 2 – 4.5; n = 16; open circle/dotted line), advanced MS participants (EDSS 5 – 7; n = 8; cross/dotted line) and controls (n = 20; closed circle, solid line). *p < 0.05; **p< 0.01; ***p<0.001 (Two-Way ANOVA, Tukey's multiple comparisons post hoc test)

4.3.9 Multivariate analysis on factors influencing performance in the spatialised sentences in babble task

To explain the impact of disease severity on spatial release of sentence discrimination from multi-talker babble, we built a model to incorporate theoretically important variables. Spatial separation in azimuth (º), disease severity, age (years), hearing sensitivity (dB HL), and trial order (1 to 30) were fixed effects included in the regression model. To test our main hypothesis that pwMS do not achieve the same degree of spatial release from masking as controls, interactions between spatial separation and mild, moderate, and severe MS groups (with controls as the reference group) were included as fixed effects. Interaction terms, viz., spatial separation*age and spatial separation*hearing sensitivity, were also theoretically important to include, as their influence on spatial release from masking effects has been previously demonstrated (268). Five theoretical regression models (**Table 3**) were generated using MATLAB Statistic Toolbox Release 2019b and compared to the constrained model to determine the difference in Akaike's Information Criterion (AIC:ΔAIC). The model with the lowest AIC was used to select the final model. Tests of fixed effects were also confirmed with likelihood ratio (LRT) tests to compare the constrained model with nested models.

The addition of interaction effects of age, hearing sensitivity and trial order with spatial separation did not significantly improve the constrained model. The model with the lowest AIC value included the addition of an interaction effect between disability group and spatial separation as a fixed effect (model 4 in bold in **Table 3**). Therefore, the constrained model was rejected in favour of the final model 4.

Parameter estimates of fixed effects are listed in **Table 4** along with the t statistic, degrees of freedom and p-values for each fixed effect to test the null hypothesis that the coefficient equals zero. The OR and confidence intervals (CI,95%) are also included to quantify the magnitude of the association between the fixed effect and the outcome.

Beta coefficients for hearing sensitivity, age and their respective interactions with spatial separation were not significantly different from 0 ($p< 0.05$), thereby suggesting that they did not influence sentence discrimination in babble (0 = incorrect; 1 = correct). Fixed effects of spatial separation, trial order, spatial separation*moderate MS and spatial separation*advanced MS were significantly associated with sentence discrimination in the SSiN task ($p<0.05$). To quantify their associations, the ORs were interpreted. When all other factors are constant, a one-unit increase in spatial separation increased the odds of discriminating the correct coordinate for controls by 5%, but only 3% for moderate and advanced MS listeners (2% lower compared to their healthy counterparts; $(1- e^{-0.02})$). There was also a positive association of trial order on SSiN performance as the odds of correctly identifying the target coordinate was 1.02 (95% CI, 1.01-1.03) times greater compared to the previous trial.

Table 3. Comparisons of fixed effects combinations in a generalized linear mixed effects model (with a logit link function) used to predict correct sentence discrimination on each trial

Constrained model (nested model):	***AIC***	***ΔAIC***	$\mathbf{X^2}$	***Δdf***	***p***
(1) Trial order + Separation (º) + Age + Hearing sensitivity + Disability group (1|participant)	2247.7	0			
Additional interaction terms					
(2) Age*Separation (º)	2249.7	+2	3.37	1	0.97
(3) Hearing sensitivity*Separation (º)	2249.6	+1.9	1.77	3	0.62
(4) Disability group*Separation (º)	**2244.2**	**-3.5**	**9.48**	**3**	**0.02**
(5) Trial order*Separation (º)	2249.5	+1.8	0.15	1	0.70

Based on 1980 observations (66 participants)
The estimate of the variability of the random effects ($\boldsymbol{\sigma_b^2}$) for all models = 0.40.
AIC = Akaike's Information Criterion
The last three columns show the chi squared statistic (X^2), difference in the degrees of freedom and p value from the likelihood ratio (LRT) tests

Table 4. Parameter estimates of fixed effects of the final generalized linear mixed effects model (with a logit link function) used to predict correct sentence discrimination from babble on each trial

						95% C.I for e^{β}	
	Estimate				OR		
Name	**(β)**	**SE β**	**tStat**	***P***	**(e^{β})**	**Upper**	**Upper**
Intercept + $N(0, \sigma_b^2)$	0.66	0.45	1.46	0.14	1.93	0.80	4.69
Trial order	0.02	0.01	3.42	**0.00**	1.02	1.01	1.03
Separation (°)	0.05	0.01	3.05	**0.00**	1.05	1.02	1.08
Hearing sensitivity (dB HL)	-0.01	0.02	-0.69	0.49	0.99	0.95	1.03
Age (years)	-0.01	0.01	-0.79	0.43	0.99	0.97	1.01
Disability group (reference = controls)							
Mild (EDSS 0-1.5)	-0.34	0.22	-1.57	0.12	0.71	0.46	1.09
Moderate (EDSS 2-4.5)	-0.29	0.24	-1.19	0.23	0.75	0.47	1.20
Advanced (EDSS 5-7)	-1.20	0.31	-3.92	**0.00**	0.30	0.16	0.55
Interaction terms (with separation (°))							
Separation x Hearing sensitivity	0.00	0.00	0.44	0.66	1.00	1.00	1.00
Separation x Age	0.00	0.00	-0.07	0.94	1.00	1.00	1.00
Separation x Mild MS	-0.01	0.01	-1.85	0.06	0.99	0.97	1.00
Separation x Moderate MS	-0.02	0.01	-2.90	**0.00**	0.98	0.96	0.99
Separation x Advanced MS	-0.02	0.01	-2.27	**0.02**	0.98	0.96	0.99

Outcome: 0 = incorrect and 1 = correct sentence
Degrees of freedom = 1967; Based on 1980 observations (N = 66 participants; one participant was omitted as age was missing)
The estimate of the variability of the random-intercept effect ($\boldsymbol{\sigma_b^2}$) = 0.39
Ordinary R^2 = 0.19; Adjusted R^2 = 0.18
SE = standard error
OR = Odds ratio
C.I = confidence interval
dB HL = Decibels hearing level
EDSS = Expanded Disability Status Scale

4.4 Discussion

Understanding speech is an important tool in forging and maintaining human relationships, and it often occurs in social, learning and work environments which contain other competing speech and non-linguistic noise originating from many directions respective to the listener. This is the first study to investigate how pwMS discriminate spatialised target speech from a competing noise source or single talker. MS listeners still benefited from spatial cues, as discrimination accuracy improved when target speech was separated from the masker by 20º and 50º azimuth compared to 0º (co-localized), however, pwMS with moderate (EDSS 2-4.5) and advanced (EDSS 5-7) disability did not achieve SRM to the same extent as controls in the SSiN task; as a one-unit increase in spatial separation increased the odds of discriminating the correct coordinate for controls by 5%, but only 3% for moderate and advanced MS listeners. Mildly affected pwMS (EDSS ≤ 1.5) achieved similar SRM to controls in the SSiN task but displayed less SRM in the CRM task where the masker was a single competing talker of the same sex. Specifically, a one-unit increase in spatial separation in the CRM task increased the odds of discriminating the correct coordinate for controls by 5%, but only 4% for mildly affected pwMS. This is the first study to describe auditory spatial processing deficits in MS, and it may partially explain the difficulties in speech understanding within complex noisy environments experienced by this clinical population.

In what appears to be a contradicting finding, the CRM sensitively detected an SRM deficit in the mild MS group, whilst the SSiN did not. The contrast in results between the two tasks is likely to reflect the different parameters used in each test, as the magnitude of SRM depends on a whole range of factors such as the spatial relationship between target and masker (symmetrical vs asymmetrical configuration)(271), the various parameters of interfering sounds (i.e. number of sounds (258), the fixed SNR etc.) and the similarity between the target and masker (272). The CRM task utilizes same-sex speakers and competing phrases with similar syntactic structure, and thereby has higher similarity between target and masker acoustic characteristics than the BKB sentences and multi-talker babble in the SSiN task. SRM is most effective when the acoustic properties of the target speech and masker are high in similarity (i.e, competing male voices) as spatial cues become the most salient cue available for parsing sound elements (251, 267, 273). This type of perceptual masking effect is termed ‘informational’, and goes beyond ‘energetic’ masking that occurs at the periphery (267, 272). Arbogast et al. (2002) reported a spatial advantage of 18dB when the masker was mainly informational, compared to a smaller release of 7dB in normal-hearing listeners when the same band noise was mainly energetic (267). Although the SSiN eight-talker babble contains spectral properties and spectro-temporal variability similar to the signal, it also contained differences in onset/offset time (babble played consistently throughout) and no semantic interference (the babble contained nonsense words), thereby, the SSiN contained more readily available segregation cues that the listener could use in addition to spatial cues- perhaps even resulting in the spatial cues being relatively ineffective. This may explain why the CRM was more sensitive in detecting an SRM deficit in the mild MS group, whilst the SSiN did not.

The ability to use spatial cues to increase speech intelligibility in the face of competing sound sources depends on the integrity of the binaural auditory system, as binaural cues are only useful when encoded with fidelity (251).

MS demyelinating lesions have a predilection for brainstem sites (260); the location where binaural interaction initially occurs to interpret ITD and ILDs (261) and for that reason, pwMS are more likely to have compromised spatial processing and reduced SRM within noisy speech processing environments. An ITD discrimination task was employed in this study to evaluate brainstem integrity for processing low-level cues, and results indicated that ITD discrimination by mildly affected pwMS (EDSS ≤ 1.5) was similar to controls on a group level. Our results contradict previous findings in another study that found significant group differences between MS listeners (EDSS scores of 0 and 1) and controls in a sound lateralization task involving ITD processing (248). This contrast in findings is likely to be driven by methodological differences, as our study compared mean ITDs +/- SEM correctly identified at six different ITD magnitudes, whilst previous investigators (248) determined a threshold ITD; a singular value to represent ITD discrimination sensitivity within 10 μs increments. It's possible that our small cohort of mildly affected pwMS did not have any ITD discrimination deficits, although, it's important to note that we considered ITD discrimination on a coarser scale (1000, 500, 250, 100, 50, 10 μs) than previous work. A slight divergence in performance between controls and mildly affected pwMS observed in **Figure 2** from 50 μs to 10 μs may be indicative of subtle differences in ITD detection that did not reach statistical significance. Nevertheless, our low-level processing task required the participant to simply detect the presence of an ITD – but not use it for segregation and across-time streaming of a target and distractor. Natural listening conditions contain a whole set of acoustic cues in addition to ITDs that can be used for grouping, including harmonicity, common onset/offset, common modulation and ILDs (274). Investigations on the role of segregation cues have concluded that ITDs present on their own are sufficient for the perceptual localisation of sounds in complex mixtures, but are not enough for the listener to perceive sound source content if no other grouping cues promoted proper source separation (275). In fact, studies of spatial unmasking during auditory scene analysis have reported that spatial cues have a weak influence in grouping simultaneous sound elements, but have a strong influence in determining which sound elements belong to which sound source at longer timescales (274). This is consistent with the CRM error analysis findings; that the majority of errors for both pwMS and controls were the coordinates (colour/number) that were from the masker phrase. This demonstrates that all listeners were able to separate target elements from each other (segregate) but made errors when attributing (or grouping) the coordinates to the correct callsign. Therefore, although ITD processing is a useful binaural cue that contributes to SRM- as confirmed by the significant predictor of ITD discrimination performance in the mixed-model analysis for SRM in 50º separation- ITD discrimination alone does not indicate how target intelligibility is achieved in the CRM.

Higher-order processes such as working memory, attention, emotion and executive function play a role in SiN processing and aid the listener to attend to relevant signal, ignore distracting irrelevant stimuli, use contextual cues to fill in words, and remember the words spoken from start to finish (210). Although cognitive influences are well established in the speech processing in noise literature (for recent reviews, refer to Peele, (2018) and Dryden et. al 2017)), comparatively less is known about the engagement of cognitive processes in spatially separated conditions. Zekveld et al. (2014), applied pupillometry to assess how informational masking from various speakers (same sex vs. different) and their location (co-located vs. separated) impacted the cognitive processing load of normal-hearing

listeners in a speech recognition task. Although both sex and location differences between target and masker facilitated speech intelligibility, the cognitive processing load was lower only when the sex of the speakers was different – but not when speakers were spatially separated (276). Lower cognitive processing load for different sex speakers is expected as a whole set of acoustic cues that can be used for grouping; for example, when female speech maskers are used for male target speech, voice-related pitch cues discriminating between the two talkers are more salient than if the masker was the same sex (276). In contrast, no difference in cognitive load when speech was spatially separate was unexpected as spatially separated speech should make it easier to segregate the target from masker, thereby placing less of a cognitive load on the listener. In contrast to this, a later study by Xia et al. (2015) utilized a different indicator of cognitive load (visual tracking) and found that spatial separation yielded a lower cognitive load (277). The contradictory findings have been attributed to experimental differences in SNRs used, whereby spatial separation of speech presented at intermediate SNRs reduced the cognitive load but adverse SNRs did not (278). The relationship between speech intelligibility, spatial separation and cognitive load is yet to be fully understood, however, cognitive abilities are a factor which must be considered in the interpretation of SRM deficits in MS – particularly as it is well recognized that MS pathology can limit cognitive resources of pwMS (57). Cognitive impairment is now considered a primary deficit affecting 40-70% of pwMS, manifesting at all disease stages, including onset (7, 236). In this study, cognition was evaluated in mildly affected pwMS through the use of the PASAT, a widely used cognitive assessment that demands speeded perceptual-motor processing, attention, working memory, and mental arithmetic, and places a heavy load on executive control processes (220). Given such a number of complex processes, reduced PASAT scores provide coarse specificity regarding cognitive operations that are impaired (279), however, it significantly predicted CRM performance in mildly affected pwMS at 0^{o}, 20^{o} and 50^{o} separation in azimuth, suggesting that cognitive deficits in pwMS may manifest as poor SiN perception.

Attention is a conscious and active process which has a robust top-down effect on most of the auditory pathway processes (280). Posner and Peterson (1990) proposed that attention can be divided into three interconnected subsystems: (a) alerting, (b) orientating and (c) executive (281). In the context of SiN processing, alerting allows the listener to concentrate on the target speech in the presence of distracting noise, orienting refers to ability to attend to speech coming from a location in space and executive attention allows the inhibition of distracting and irrelevant sensory input streams in favour of the target voice (281, 282). Of particular relevance to this study, is the process of orientating attention to prioritize speech coming from a specific location in space, and reduce the interference caused by a masker at a different location (281, 282). At a colocalized condition (0^{o}), it's difficult for the listener to separate the signal from the informational portion of the masker because there aren't many salient differences between the two – in fact, the CRM was a very difficult task for all participants; anecdotally, two control listeners expressed that they found the task quite difficult and hard to tolerate and were unable to complete it. Hence, no differences in performances were found between the controls and mildly affected MS group at the very difficult colocalized condition. It's possible that when the masker is moved away from the target, the bottom up salient spatial cues guide the top-down attentional processes required to orient auditory attention to the target speaker. The spatial binaural cue produced a greater

listening advantage in controls than the cognitively impaired MS group, as evident by deficits at the 50º spatial separation but not at the colocalized position in the CRM.

Up until now, not much was known about how pwMS navigated complex and dynamic acoustic environments with respect to the spatial configurations of sound sources. This is the first study to describe deficits in spatial release from masking within MS, thereby exposing an important aspect of successful speech processing in everyday life previously overlooked.

Chapter Five: An Exploratory Study on the Relationship Between Brain Atrophy and Processing Degraded Speech Abilities in Multiple Sclerosis (Preliminary Findings)

Preface

The work presented within the previous chapters of this thesis support the notion that speech-in-noise (SiN) tasks could be useful for disease surveillance in multiple sclerosis (MS). **Chapter Two** demonstrated that signal-to-noise ratio may be a useful biomarker for disease burden, whilst **Chapter Three** confirmed that SiN tasks sensitively detect subtle MS deficits occurring in the early stages of insidious progression that often go overlooked by current "gold standards". Furthermore, SiN processing abilities in both non-spatial and spatial conditions (**Chapter Four**) reflect cognitive impairment as measured by standard neuropsychological assessments. The next step to validating the clinical utility of SiN tasks for measuring sensory-cognitive changes in MS is to explore the pathological underpinnings of SiN deficits in MS using Magnetic Resonance Imaging (MRI) techniques. Brain volume loss is a widely recognized pathological feature of MS that is an indicator of neurodegeneration and a sound predictor of cognitive impairment, therefore, in this chapter I investigated the association between our SiN tasks and volumetric measurements of several brain structures posited to be implicated in processing speech in the presence of noise. Understanding pathological associations will strengthen the interpretation of MS-deficits in our SiN tasks, which I have previously posited to represent sensory and cognitive deficits.

The study presented is highly exploratory, therefore these preliminary findings have not been prepared as a manuscript for submission.

5.1 Introduction

In the present study, I examined the relationship between cortical anatomy and the ability to perceive speech-in-noise (SiN) in people with multiple sclerosis (pwMS). The work presented within the previous chapters of this thesis, described a deficit in perceiving SiN in normal hearing pwMS that is more severe in individuals with greater disability. **Chapter Two** demonstrated that psychometric curves, which model the relationship between signal-to-noise ratio (SNR) (i.e. the volume of the target speech in relation to the background noise) and verbal recall of sentences, were comparable in slope gradient (sentences/dB) but were shifted to higher (easier) SNRs (dB) in MS listeners compared to controls. The direction of shift suggested that MS listeners needed higher SNRs to achieve normal speech discrimination accuracy, and the magnitude of the shift systematically increased with greater disease severity, providing preliminary evidence that SNR may be a useful biomarker for monitoring disease burden. **Chapter Three** demonstrated that SiN tasks sensitively detect MS-group deficits in the early stages of disease, prior to the development of overt physical symptoms. This may be informative for prognostication. The next step in determining the clinical utility of SiN tasks for disease surveillance, is to explore the pathological underpinnings of SiN deficits in MS using Magnetic Resonance Imaging (MRI) techniques. This will strengthen the interpretation of MS-deficits in our SiN tasks, which I have previously posited to represent sensory and cognitive deficits in MS individuals.

Speech communication rarely occurs in silence, as most real-world conditions contain some degree of noise. A body of evidence suggests that perceiving speech in the presence of noise requires more effortful listening than clear speech, due to greater cognitive resources such as attention to focus on target speech whilst simultaneously ignoring distracting noise and working memory to restore impoverished low-level sensory representations through context and prior lexical knowledge (201). Throughout this thesis, we have focused on two types of noise interference in particular. Although discussed in earlier chapters, a brief overview of the properties of the acoustic maskers is important to discuss here. Speech-weighted noise utilizes "energetic masking" to diminish audibility of a target from interference of shared spectro-temporal acoustic signals in the lower levels of the auditory system (210), whilst multi-talker babble involves energetic interference but also "informational masking" that produces high-level attention competition effects due to confusability of similar target and masker (212, 213). As postulated in **Chapter Three**, the multi-babble masker is a particularly potent masker for early pwMS due to its speech-like attentionally-demanding, salient acoustic characteristics resulting in greater speech discrimination impairment in multi-talker babble, but not in speech-weighted noise. Furthermore, correlations between standard neuropsychological tasks and speech discrimination in only babble- but not speech-weighted noise- in early pwMS described in **Chapter Three** support the view that babble engages greater cognitive resources to extract the target speech. Poor SiN perception in pwMS may be a manifestation of cognitive deficits, which affects up to 70% of pwMS (7, 188). Cognitive domains typically affected in MS are executive function, attention, memory, and information processing speed (57, 58).

Functional neuroimaging methods have been widely used to explore the interplay between sensory and cognitive processes that are engaged to support the perception of degraded speech. As summarised in a review by

Scott and McGettigan (2013), neuroimaging studies of brain regions involved in processing degraded speech demonstrate various activation patterns of cortical and subcortical regions influenced by different linguistic stimuli, SNRs and masker types (283). Generally, more widespread prefrontal and parietal activation outside the primary auditory cortex is associated with SiN compared to clear speech (283, 284). Prefrontal areas associated with attention and working memory are anatomically connected to auditory belt and parabelt regions, and are therefore able to modulate early auditory processing (285). Indeed this pattern of activation is commonly seen in cognitive tasks requiring explicit attentional mechanisms (286), across modalities (287, 288). Activation of premotor areas, in particular the left premotor region, during SiN processing may suggest that the mapping of perceived speech sounds onto articulatory motor representations of speech may also be an important component of understanding speech in adverse listening conditions (289, 290). Greater activation of networks beyond the temporal lobe has also been described by Wong et al. (2009) through the use of functional magnetic resonance imaging (fMRI) in older participants with expected age-related sensory and cognitive changes. Behaviourally, when words were presented in silence or in multi-talker babble at a high SNR, older and younger listeners performed similarly, however, when speech was presented in multi-talker babble at a low, more difficult SNR, older listeners performed less accurately. fMRI showed increased cortical activity in working memory and attention-related cortical areas (prefrontal and precuneus regions) in older participants, particularly at a low SNR (291). Age-related under activation detected via brain imaging is widely accepted to be an impairment due to age-related atrophy or structural change, whilst the neurophysiological characteristics of overactivation are more contentious in confirming whether overactivation is beneficial or detrimental to cognitive function (292). Wong et al.'s 2009 study confirmed the *decline-compensation hypothesis* relevant to the auditory processing domain; a theory which posits that overactivation (relative to younger counterparts) in prefrontal areas is associated with better cognitive performance in the elderly. Specifically, findings confirmed an under activation in the sensory-encoding areas of the bilateral superior temporal cortex and overactivation in prefrontal areas associated with attention and working memory in the older group. Greater accuracy in word processing in noise was achieved in older participants with increased activation in frontal and posterior parietal networks, suggesting a compensatory role in aiding older participants with age-related sensory decline (291).

Taken together, behavioural and neurophysiological studies suggest that SiN perception requires both sensation and cognition (201, 283). In the present study, we investigated the possible link between anatomical characteristics of brain regions and SiN perception abilities in pwMS. PwMS exhibit pronounced grey matter atrophy in all brain regions, including the basal ganglia, cortex, cerebellum and brainstem (293), at all stages of the disease. Several mechanisms may underlie brain atrophy, including extensive axonal transection and demyelination leading to retrograde neurodegeneration, cortical inflammatory pathology, mitochondrial failure, and iron deposition (294, 295). Volume loss is a widely recognized pathological feature of MS indicating neurodegeneration, and is a sound predictor of cognitive impairment (296). Further, enlargement of the third ventricle and atrophy of the thalamus which mediates cognitive function via cortical and subcortical pathways, are also clinically relevant biomarkers of neurodegeneration (296, 297). Conversely, correlations between cognitive status and WM lesion load on T2-weighted MRI are modest (298). We hypothesized that neurodegeneration in prefrontal and parietal areas involved in high-order cognitive

processes, may underlie SiN processing deficits in MS. We quantified neurodegeneration in pwMS as regional atrophy (i.e. volume loss) on clinical MRI scans and determined the association between our SiN tasks and volumetric measures of several key neuroanatomical regions, namely, the thalamus and frontal, parietal and temporal cortices. The Paced Auditory Serial Addition Test (PASAT), a standardized neuropsychological assessment used to evaluate working memory, mental arithmetic, and information speed, that similarly requires the registration of auditory input, the deployment of central cognitive resources, and the generation of a verbal response (202) was also evaluated in pwMS. It was hypothesized that the poorer SiN and PASAT performance by a subset of pwMS described in **Chapter Three** would correlate with reduced brain volumes in the thalamus, prefrontal, parietal and temporal regions.

5.2 Methods

Please refer to **Chapters Two, Three and Four** for details regarding patient recruitment and ethical procedures. To avoid repetition, only methods pertinent to the aims of this chapter have been included below**.**

5.2.1 Participants

MS patients with a relapsing-remitting time-course were included in this study, with diagnosis based on McDonald's criteria (29). All pwMS were independently mobile, with little to no disability (Kurtzke Expanded Disability Status Scale Scores (EDSS)(171) of ≤ 1) and continued to take all prescribed medications. No MS participant had experienced a recent (within 30 days) relapses and/or steroid administration.

5.2.2 Neuropsychological testing

Only pwMS performed the PASAT (220).

5.2.3 Magnetic Resonance Imaging

MRI scans were performed as part of each patient's regular clinical monitoring. Images were obtained using a 3-Tesla MRI system (Trio TIM, Siemens, Erlangen, Germany) with a 12-channel head coil and a slice thickness of 5mm. A 3D whole brain T1-weighted magnetisation-prepared 2 rapid gradient echoes (MP2RAGE) sequence was acquired for volumetric analysis. This sequence allows efficient 3D mapping of T1 scans (299). Volumes of cortical and subcortical structures were calculated using the MorphoBox algorithm, which automatically estimates normative ranges adjusted for head size, sex, and age. Brain volumes are expressed as a percentage of total intracranial volume. Based on structures involved in models of SiN discrimination, brain volumes were calculated bilaterally for the thalamus, frontal, parietal and temporal lobes.

5.2.4 Audiometry

Hearing sensitivity was determined for all participants. For methods, refer to Audiometry sections outlined in **Chapters Two, Three and Four**.

5.2.5 SiN Discrimination Tasks

The general procedures and stimuli for the sentences in speech-weighted noise (SWN) and multi-talker babble noise (BN) task have been detailed previously in **Chapters Two and Three**.

5.2.6 Statistical analysis

Statistical analyses were performed using IBM SPSS v26. Boltzmann sigmoidal functions using Graphpad Prism 8 were fitted to obtain psychometric curves as a function of SNR for individual participants in each SiN task (see **Figure 1**). The top and bottom of the functions were constrained to 10 and 0 sentences correct, respectively. Slope

(sentences/dB) and midpoint data (dB) from the curves were extracted and compared using unpaired t-tests. Pearson product-moment correlations were used to analyse the relationship between MRI measures and SiN tasks.

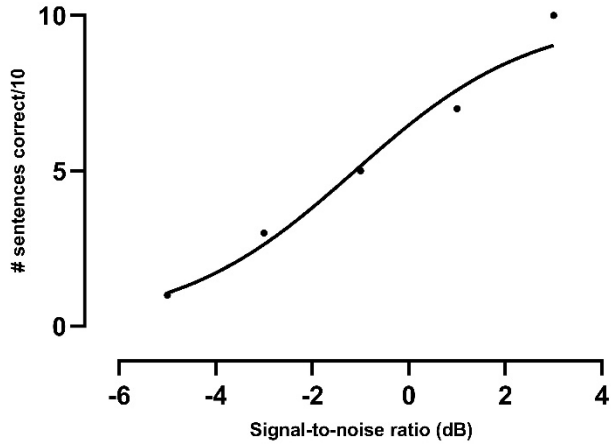


Figure 1. A visual example of a Boltzmann sigmoidal function (black line) fitted to a participant's scores out of ten at five signal-to-noise ratios (SNR) (black circles). The slope (sentences/dB) and midpoint (dB) of the curve was extracted for statistical analysis between control and multiple sclerosis groups.

5.3 Results

5.3.1 Participant demographics, characteristics, and audiometric hearing status

20 controls and 17 MS participants were evaluated in this cross-sectional study. MS participants in this study are a subset of those described in **Chapter Three**. For participant details, refer to **Table 1**. Participants evaluated in the current study had bilaterally normal hearing between 250 – 4000 Hz; except for 5% of participants from each group who had small hearing losses (of 5–10dB) at the higher frequencies of 6000 and 8000 Hz in one ear only. Potential confounders such as age, and the four-frequency pure tone average (the average hearing threshold across 500, 1000, 2000 and 4000Hz in each person; a standard audiometric measured to index a person's hearing) (dB HL) in the left and right ears were not statistically different between controls and pwMS ($p>0.05$).

Table 1. Participant characteristics

		Control	MS	t	df	ŋ²
Demographics	**Number of participants**	20	17			
	Sex F(M)	17 (3)	15 (2)			
	Age, (yrs)					
	Mean (SD)	45.85 (10.82)	45.94 (9.11)	0.03	35	0.00
	Range	28 - 60	28 – 63			
Audiometry	**Pure tone average (dB HL)**					
	Left (Mean, SD)	13.12 (5.90)	13.40 (4.46)	0.06	35	0.00
	Range	-1.25 to 18.75	5 to 18.75			
	Right (Mean, SD)	12.06 (5.13)	13.34 (5.48)	0.93	35	0.02
	Range	2.5 to 23.75	2.5 to 21.25			
Disease Characteristics	**Disease duration, (yrs)**	-				
	Mean (SD)	-	9.29 (5.27)			
	Range	-	1 - 18			
	EDSS	-				
	Median	-	0			
	Range	-	0 - 1			
	Disease modifying therapy (%)	NA	94.11%			

df = degrees of freedom
F = female; M = male
SD = Standard deviation
dB HL = decibels hearing level
EDSS = Expanded Disability Status Scale Score
NA = not applicable

5.3.2 Discrimination of sentences in noise

To quantify MS effects on sentence discrimination, Boltzmann sigmoidal functions were fitted to each participant's discrimination curve. Measures of goodness of fit were strong for each group (R^2 always > 0.9). From each such psychometric curve, the slope (sentences/dB) and midpoint (SNR (dB)) were extracted, and a unpaired t test revealed no significant difference in mean slopes of the psychometric functions between control and MS listeners for speech discrimination in SWN [$t(35) = 0.41$, $p = 0.68$, $\eta^2 = 0.004$] or in BN [$t(35) = 0.12$, $p = 0.90$, $\eta^2 = 0.0004$].

The midpoint SNRs of the curves are graphed in **Figure 2A and 2B** for sentence discrimination in SWN and BN, respectively. An unpaired t-test indicated that the midpoint SNRs of control and MS psychometric functions for sentence discrimination in SWN were not significantly different [$t(35) = 0.91$, $p = 0.37$, $\eta^2 = 0.02$]. In contrast, midpoint SNRs of control and MS psychometric functions for speech discrimination in BN were significantly different [$t(35) = 3.10$, $p = 0.004$, $\eta^2 = 0.21$]. This is consistent with the findings in **Chapter Three** and confirm that pwMS were impaired in SiN processing, informing the relationship with neural changes in MS.

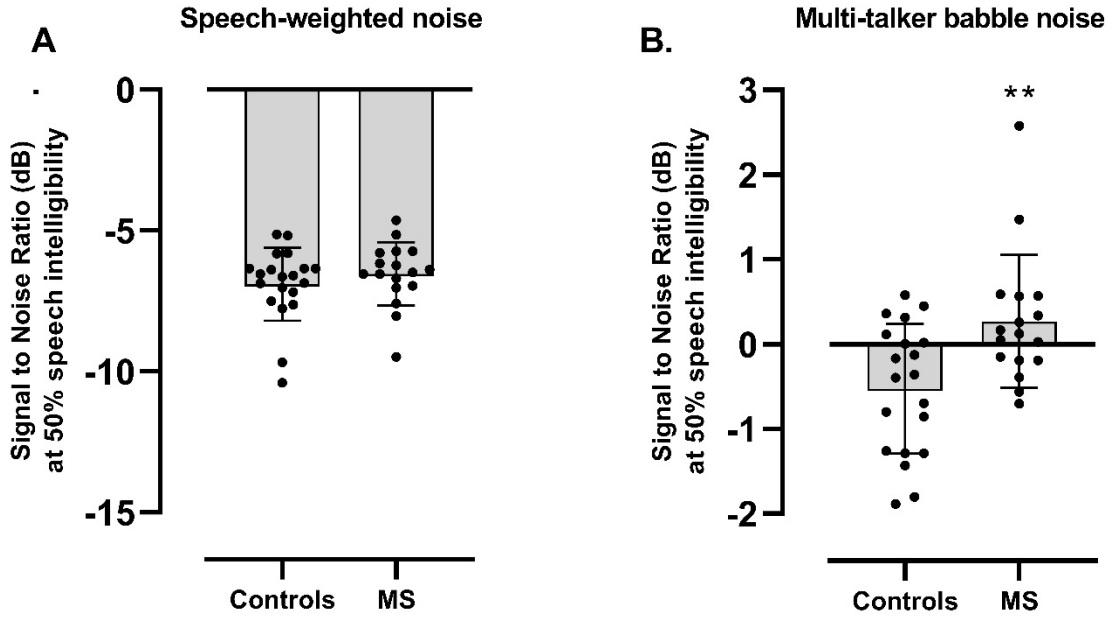


Figure 2. People with MS had greater difficulty with sentence discrimination in an attention-demanding babble masker but not in energetic speech-weighted noise. Mean signal-to-noise ratio (SNR) ± SD (dB) at 50% speech discrimination in speech-weighted noise **(A)** and multi-talker babble **(B)** for control (n = 20) and MS (n = 17) listeners. SNRs at 50% discrimination in multi-talker babble were statistically higher for MS listeners compared to controls. **p < 0.01; compared to controls. Unpaired t-test.

5.3.3 Magnetic Resonance Imaging analysis

Brain volumes were expressed as a percentage of total intracranial volume. Average percentage volumes of the thalamus, frontal WM, frontal GM, parietal WM, parietal GM, temporal WM and temporal GM , are presented in **Table 2**, as well as the number of MS patients with significantly reduced volumes relative to controls.

Table 2. Means, standard deviation and number of MS participants with significantly reduced MRI volume for various brain structures

Region	Mean^	SD	Number of patients with significantly reduced volume*
Grey Matter	46.74	2.14	0
Third ventricle	0.11	0.03	0
Thalamus	0.91	0.06	11
Frontal white matter	10.64	0.80	1
Frontal grey matter	13.28	0.90	1
Parietal white matter	6.53	1.29	2
Parietal grey matter	9.48	0.96	0
Temporal white matter	4.23	0.47	1
Temporal grey matter	8.96	0.45	0

MRI = magnetic resonance imaging

SD = Standard deviation

^ Measured as a percentage of total intracranial volume

*Compared to normative data in the MorphoBox Algorithm

Correlations between volumetric measures and performance on the SiN tasks are presented in **Table 3**. Correlational analyses revealed a significant negative relationship between SNR (dB) at 50% speech intelligibility in BN and temporal lobe WM volumes (%), ($r(17) = -0.59$, $p = 0.01$) (presented in **Figure 3A**), and a significant negative relationship between SNR (dB) at 50% speech intelligibility in BN and PASAT scores (%), $r(16) = -0.60$, $p = 0.01$. No other correlations were significant. A non-significant correlation between the SNR (dB) at 50% speech intelligibility in BN and frontal lobe WM volume (%) is presented in **Figure 3B,** to visually show that SiN measures did not correlate to frontal volume as hypothesized.

Table 3. Pearson correlation coefficients (r) to describe the relationship between MRI volumetric measures, neuropsychological assessment, and speech-in-noise tasks

Volumetric measures of global brain regions		
	GM	**Third Ventricle**
Sentences in speech-weighted noise	0.47	-0.29
Sentences in multi-talker babble	0.44	-0.15
PASAT	0.03	-0.04

Volumetric measures of subcortical and cortical regions							
	Thalamus	**Frontal GM**	**Frontal WM**	**Parietal GM**	**Parietal WM**	**Temporal GM**	**Temporal WM**
Sentences in speech-weighted noise	-0.19	0.20	0.09	0.09	0.03	0.46	-0.14
Sentences in multi-talker babble	-0.06	-0.03	-0.05	0.13	0.06	-0.04	**-0.59****
PASAT	-0.14	0.03	-0.01	0.38	0.03	0.32	0.31

Neuropsychological assessment	
	PASAT
Sentences in speech-weighted noise	-0.12
Sentences in multi-talker babble	**-0.60****

MRI = magnetic resonance imaging
GM = grey matter
WM = wihte matter;
PASAT = Paced Auditory Serial Addition Test
**p = 0.01

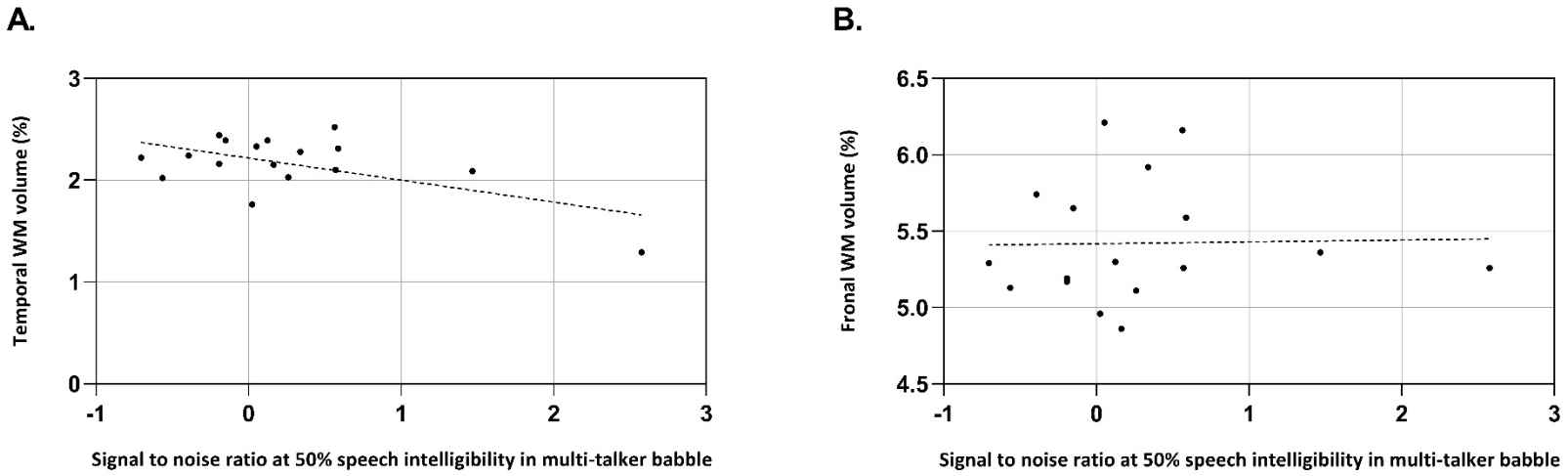


Figure 3. Poorer speech-in-noise (SiN) performance was associated with smaller temporal white matter (WM) volume (%) only. A significant correlation between the signal-to-noise ratio (dB) at 50% speech intelligibility in multi-talker babble and temporal white matter volume for people with multiple sclerosis (pwMS) **(A)**. Non-significant correlation between the signal-to-noise ratio (dB) at 50% speech intelligibility in multi-talker babble and frontal white matter volume **(B)**.

5.4 Discussion

To support my proposal that SiN tasks are a sensitive measure of sensory and cognitive changes in pwMS, I sought to clarify the relationship between SiN task performance and volume loss as an indicator of neurodegeneration (300). Consistent with our previous findings, pwMS (with minimal disability; EDSS ≤ 1) performed more poorly on the SiN task than controls where sentences were presented in multi-talker babble, but not speech-weighted noise. A significant relationship between speech in babble performance and PASAT scores was revealed, with lower PASAT scores associated with poorer SiN performance (represented as higher SNRs at 50% intelligibility). There was also a significant relationship between temporal WM volume and speech discrimination in babble performance, with poorer SiN performance associated with smaller temporal WM volume (%). There were no other significant associations between SiN performance and other brain region volumes. It should be noted that the association between speech discrimination accuracy in babble and temporal WM volume is weak and largely driven by a single participant who had an atypically high SNR at 50% discrimination accuracy. Therefore, it's possible that this association may not necessarily be present in the larger clinical population and tentative interpretations of this association remain subject to further confirmation until results are reproduced in a larger sample.

The processing of audible speech in silence is associated with subcortical and cortical processing, with the primary auditory cortex (PAC), located bilaterally on the superior temporal gyrus (STG) of the dorsolateral temporal lobe, receiving input from the auditory thalamus (301, 302). The earliest stages of cortical processing of speech involve spectro-temporal analyses (i.e. incorporating information about the frequency and timing of speech) (303) predominantly within the left superior temporal lobe (STG) and superior temporal sulcus (STS). The precise spectro-temporal features are arranged into phonemes in the secondary auditory areas located in the middle to posterior portions of the STS, and then those phonemes are arranged into words(304). The well-accepted dual-stream model of language by Hickok and Poeppel (2007) describe two predominant processing pathways in the auditory cortex: the ventral and dorsal pathways. There is wide consensus that word recognition occurs in the ventral auditory stream (303) and phonological representations are mapped onto articulatory motor representations in the dorsal auditory stream, that projects posteriorly and dorsally (302). Determining the meaning of a word sequence, based on syntactic and semantic properties, involves hierarchical high-level networks of temporal, parietal and frontal areas (304, 305).

Beyond the temporal lobe, there were no other associations between SiN performance and other cognitive-associated brain regions as hypothesized. Wong et al. (2011) found that a larger/thicker prefrontal cortex is related to better SiN processing in background babble in older adults with age-related brain atrophy (306), in contrast to this, we did not find the same link between frontal lobe volume and SiN processing in pwMS. Thalamic volume has also been strongly linked to neuropsychological test performance in pwMS, indicating an association between cognition and atrophy in the thalamus (297, 307), however, this is another association we did not find in this small study of participants with minimal impairment. The lack of correlations between SiN performance in speech-weighted noise and brain volumes in pwMS in our study is consistent with lack of impairment here. As for the babble task, it's possible that a lack of correlation outside of the temporal lobes may have reflected a lower cognitive demand required for the

listener to parse the competing streams in babble; however, scores from the PASAT, a widely used and well-established neuropsychological assessment of cognition, also had no correlations to any measures of brain volume in pwMS. Previous studies reveal inconsistent findings in terms of the neural correlates of PASAT performance. Studies have variously revealed correlation with atrophy of the prefrontal cortex, superior parietal cortex and right cerebellum (308), atrophy of multiple subcortical structures (309), and atrophy of no brain region (310). It has been suggested that participants completing the PASAT may use different strategies to complete the task, such as skipping every third digit to reduce the difficulty of the test to achieve higher, yet unreliable scores (311, 312). This may partially explain the lack of correlation found here. Several studies have also reported enhanced compensatory mechanisms from cortical reorganization in pwMS during the PASAT to perform within normal limits (66-68, 137, 313). Neuroplasticity in MS to achieve normal functioning is a strong contributing factor to the 'morphological-clinical gap' as morphological insights by using MRI only allow assessment of structural disease changes of the brain, whilst functional imaging may detect adaptive changes (298). This may also explain the lack of brain region volumes correlating to SiN perception abilities in pwMS in our study, therefore future studies should utilize functional neuroimaging methods to interpret SiN performance in pwMS.

Although I revealed that performance on the SiN (babble) task and PASAT were correlated, largely reflecting similar working memory and attention capabilities (220), we must acknowledge that the two tests require the engagement of slightly different cognitive processes. The manner in which auditory stimuli are presented differs in that digits in the PASAT are presented individually in silence whilst stimuli in the SiN tasks are presented simultaneously and are complex, increasing the demands on sensory processing. This may explain the correlation between temporal lobe WM volume and SiN performance, but not PASAT scores.

A limitation of this study was that it was not possible to measure the volume of discrete regions within temporal, frontal and parietal lobes. This may have provided a more detailed understanding of the neuropathological underpinnings of SiN deficits in pwMS. Future studies should parcellate these structures and evaluate MS participants across a wider spectrum of severity (as we only looked at participants EDSS scores ≤ 1), particularly since the neuropathology in more progressive forms of MS is largely driven by neurodegenerative processes rather than inflammatory processes (314). Finally, the use of a longitudinal study design is needed to confirm the interplay between SiN processing and neural changes over time in MS. Current knowledge of how MS-related structural damage in the brain affects SiN processing is preliminary and highly exploratory, therefore results should be replicated.

Chapter Six: General Discussion

Multiple sclerosis (MS) is a complex, highly heterogenous disease of the central nervous system with a largely unknown etiology, that afflicts over 2.5 million individuals worldwide (7, 315). Communication problems are commonly reported, and even a mild communication problem can have significant impacts on social relationships and the ability to gain and maintain employment (316); this is of particular concern in MS as disease onset is typically 20 to 40 year of age (i.e. the most productive years of life) (71).

Speech is the main form of communication that keeps social human beings connected, and it often takes place in the presence of ambient noise or other competing speech that can originate from various locations in three-dimensional space (180, 212). Successful speech perception in the presence of noise is dependent upon the integrity of a dynamic, integrated system involving both bottom-up (encoding incoming auditory stimuli) and top-down (cognitive) processing. Bottom-up, or stimulus-driven processes refer to detecting and automatically organizing the acoustic characteristics of an input into distinct streams based on spectro-temporal characteristics (317). This relatively primitive process plays an important role in how top-down or higher-processes such as attention interact to refine and highlight what we aim to perceive (230). A common theme in this discussion will be how MS pathology affects both of these processes, leading to deficits in understanding speech in the presence of noise.

This discussion will integrate findings across each of the chapters of this thesis and discuss their significance and implications. These findings can be summarized into two main themes:

1) The effects of MS on auditory processes
2) The utility of speech-in-noise (SiN) tasks as a clinical measure to evaluate cognitive and sensory impairments in early and mild MS

6.1 Study design aspects

Before an in-depth discussion of the observed effects, it is important to recall how patient groups were characterised. The literature on auditory deficits in people with MS (pwMS) is conflicting, generally attributed to the heterogeneity of the disease (79). Furthermore, progressive MS phenotypes such as secondary progressive (SP) MS, largely drive significant effects described in observational studies. The strength of the studies reported herein was the sub-group analysis based on stratifying patients according to their Disability Status Scale (EDSS) scores. A number of previous studies have used EDSS scores to infer cognitive impairment, despite the fact that EDSS scores largely reflect physical impairment and ambulatory status (186). The relationship between physical disability and cognitive impairment in MS is complex and conflicting (227); some studies demonstrate that EDSS scores do not predict cognitive status (318) while others describe some association between cognition and physical disability (228). However, while the EDSS

has its limitations beyond minimal evaluation of cognitive impairment - such as high subjectivity and uneven steps of progression - it is the current gold standard, used by clinicians to measure disease progression (319).

6.1.1 EDSS classification groups

MS participants in studies in **Chapters Two** and **Four** were grouped according to EDSS scores to determine how disease severity impacted SiN processing under various conditions. The classification of MS participants was based on that used in other studies in the field to classify mild (EDSS 0-1.5) (320-322) and advanced disability (EDSS 5-7) (323). Moderate disability was classified in between these two boundaries (EDSS 2-4.5). The boundaries used are also congruent with the classification scheme from MS support organisations that describe an EDSS score 0-1.5 as 'no disability' (324, 325) and EDSS 5+ as severe (324).

With respect to the demarcation between mild and moderate disability, an EDSS of 2, although representing mild disability on some scales, is a threshold where patients are aware of noticeable neurological problems. Other lines of evidence suggest that an EDSS of 2 is a boundary for impairments in cognition. After a subgroup analysis, Migliore et al. (2017) found deficits in the cognitive domains of verbal memory and executive functions in those EDSS $\leq$ 1.5; whilst another group EDSS (2 - 2.5) were additionally impaired in information processing speed and visual memory (322). With respect to treating pwMS with EDSS scores of 5 and above as being advanced MS cases, EDSS steps 1.0 - 4.5 refer to patients who are fully ambulatory (the precise step number being defined by the functional system score) and EDSS steps 5.0 – 9.5 are defined by the degree of impairment to ambulation (324, 326) – i.e., there is a distinct shift in ambulatory function between EDSS steps <5 and EDSS steps of 5 or greater. This is also seen in other markers of MS. Inflammatory markers (IgM, homocysteine levels, CRP and NLR) in MS patients are higher in patients with EDSS scores $\geq$ 5 than in MS patients with EDSS scores <5, consistent with the change in ambulatory impairment at EDSS scores of 5 or greater compared to lower scores (323).

Conversely, studies reported in **Chapter Three** focused on early disease monitoring, restricted to patients with an EDSS < 3. This cut-off was chosen because it is clear that once a pwMS obtains an EDSS score > 3, disease-modifying therapies (DMTs) are less effective (3, 327).

6.1.2 Sub-group analysis vs. single case reports

All research chapters within this thesis were cross-sectional group studies. Indeed, MS can present heterogeneity in disease progression and causes, yet, it is noteworthy that when the EDSS disability score was used to group pwMS, the effects were uniform within each cohort (mild, moderate and advanced) across the diversity of tests used (**Chapter Two and Four**). This suggests that the processes underlying speech processing and communication deficits appear to be quite uniform when indexed to the "gold standard" measure of MS disability, the EDSS. Case series or single case reports are observations on a series of individuals without a control group, and are advantageous for investigating accidental or rare phenomena, making new observations, forming new hypotheses, and providing in-depth narrative studies on a patient's perspective (328). Given that MS is a heterogenous disease, this type of qualitative data would

be highly valuable for the investigation of new therapeutics in rarer clinical phenotypes of MS; however, this type of design is limited as it provides less quantitative data, cannot sufficiently prove causal relationships and runs a risk of selection bias (329). In contrast, cohort studies allow the evaluation of the relationship of a particular factor (such as EDSS score, disease modifying treatment) and the outcome under investigation, and therefore inform about the chance of such changes in similar populations (329). Cohort study designs were ideal for initial investigations into SiN performance in MS to establish general associations between disease factors and SiN performance, however, it should be noted that grouping participant data runs the risk of underestimating or overestimating the relationship between variables, particularly with smaller sample sizes, which is problematic in a heterogenous clinical population such as MS. The ideal study design for clinical MS research will depend on the research question, quality of data and the analysis employed.

6.2 Auditory processing in MS

The two major domains of hearing are (1) peripheral - the transmission of sound waves through the auditory periphery (outer, middle and sensorineural processing within the cochlea) and (2) central - the subcortical and cortical processing for the analysis and interpretation of sound (brainstem, midbrain and auditory cortex) (330). As such this section will be structured according to the MS-related auditory impairments observed in each domain.

6.2.1 Peripheral auditory processing in MS

Peripheral hearing loss can largely be classified into two categories: a) conductive and b) sensorineural (331). As conductive hearing loss involves blockage or damage to anatomical structures in the outer and middle ear, MS-demyelinating pathology does not directly cause this type of hearing loss and cases are rarely, if at all, diagnosed (78, 332). As conductive hearing loss was not specifically measured in any of our studies, it is impossible to know if hearing loss detected by pure-tone audiometry in control and MS participants (**Chapters Two, Three, Four** and **Five**), was due to conductive, sensorineural or a mix of both types of hearing loss. Only those participants with normal bilateral thresholds (≤ 25 dB) were included in the studies to eliminate the possibility of peripheral hearing loss as a confounding factor in the interpretation results. Although we did not test enough participants to determine the prevalence of pure-tone thresholds in MS, we excluded a similar percentage of people in the control and MS groups in all studies.

Sudden sensorineural hearing loss is estimated to occur between 1 – 17% of pwMS and is defined as a hearing loss of at least 30 dB in three sequential frequencies in the audiogram occurring within three days or less (76). However, a well-controlled study reported that pure-tone thresholds are not chronically influenced by MS (81), and rare cases of hearing loss are likely to be temporary and associated with exacerbations of the disease (82, 83). For this reason, pwMS who experienced exacerbations within three months were excluded from all studies. Although individuals with hearing loss were excluded from our/my studies, future investigations in pwMS with peripheral hearing loss are recommended. Studies of age-related hearing loss indicate that long-term sensory deprivation can

accelerate cognitive decline and negatively impact psychosocial well-being; at the same time, changes in cognition might also be a direct result of using additional resources such as attention and working memory to compensate for sensory loss (333, 334). Nevertheless, the relationship between peripheral hearing loss and cognition is complex and likely to include other factors such as age (334). In a clinical cohort where cognitive decline affects up to 70% of individuals (58), the potential compounding effect of hearing loss may be particularly detrimental. This remains to be investigated in pwMS, however, sensorineural hearing loss is considered treatable by hearing aids and cochlear implants (316, 333, 334), and growing evidence supports the benefit of cochlear implantation treatment for oral communication, quality of life and neurocognitive functioning more generally (335).

6.2.2 Central auditory processing in MS: Brainstem processing

Processing in the auditory brainstem is particularly vulnerable to MS demyelination which slows conduction and causes neuronal dyssynchrony, manifesting as deficits in sound analyses involving precise neuronal timing (10, 11). The purpose of evaluating interaural timing cues (ITDs) within a mildly affected MS group in **Chapter Four** was twofold: (1) to evaluate the integrity of brainstem processes (2) and determine the extent to which impoverished timing detection within the brainstem contributes to SiN processing deficits. Exquisite timing is important for capturing fast-changing acoustic transitions, not just within speech itself (monoaural cue), but also in spatially separated conditions (binaural cue) (164).

Our preliminary findings revealed that ITD discrimination by mildly affected pwMS (EDSS $\leq$ 1.5) was similar to controls on a group level. These results contradict previous findings of a significant group difference between MS listeners (EDSS scores of 0 and 1) and controls in a sound lateralization task involving ITD processing (248). This is likely to driven by methodological differences, as our study compared mean ITDs +/- SEM correctly identified at six different ITD magnitudes, while previous investigators (248) determined a threshold ITD; a singular value at which participants could correctly localize sound within 10 μs increments. The coarse scale of ITD magnitudes measured here may have reduced sensitivity and requires clarification in further studies. Electrophysiological measures of amplitude and latencies within an evoked auditory brainstem response (ABR) could provide an objective means to confirm abnormal processing in the auditory brainstem as there would be clear comparisons of the frequency and timing components of the stimulus and response (336).

It is possible that there were, indeed, no ITD processing deficits within our small mildly impaired MS group; but caution must be taken when concluding that MS-deficits in SiN processing are not attributable to brainstem timing deficits. Our low-level processing task required the participant to simply detect the presence of a binaural timing cue embedded in a simple noise burst; while noise bursts and pure tones are informative for basic processing, they do not translate to processing behaviorally relevant sounds such as listening to speech in a noisy room (165). Speech is a complex sound consisting of rapid spectro-temporal fluctuations, rich harmonic structures, and dynamic amplitude modulations that the brainstem must faithfully encode during the segregation and streaming process of analysing an auditory scene (163-165). PwMS are unlikely to perceive this type of rapid information with the precision required

due to slowed or less coordinated neural conduction. Another important aspect to consider is the extensive efferent fibres from the cortex to the inferior colliculus within the brainstem, supporting the proposal that the auditory brainstem also receives input from higher-order processes that can modulate processing (163, 337). Since MS lesions form indiscriminately, it is possible that damage at various anatomical sites within the auditory brainstem contributed to poorer SiN perception observed in pwMS.

6.2.3 Central auditory processing in MS: Speech-in-noise processing

This thesis characterised impairments in SiN perception within this clinical cohort and sought to understand the extent to which impairments disrupted daily life. We were the first to investigate this using psychoacoustic tests which employed ecologically relevant stimuli such as multi-talker babble and open-set natural whole sentences coupled with self-reports on the Auditory Attention and Distress Questionnaire (AADQ). **Table 1** provides a summary of the SiN processing deficits we observed in our body of work which will be discussed in detail within this section.

Table 1. Summary table of speech discrimination impairments (group-level) in MS with various disability severity

Speech Discrimination Assessment	Mild (EDSS ≤ 1.5)	Moderate (EDSS 2-4.5)	Advanced (EDSS 5 – 7)
Sentences in speech-weighted noise (**Chapters Two and Three**)	No impairment	Shift of psychacoustic curve to higher SNRs; no change in slope	Shift of psychacoustic curve to higher SNRs; no change in slope
Sentences in multi-talker babble (**Chapters Two and Three**)	Shift of psychacoustic curve to higher SNRs; no change in slope	Shift of psychacoustic curve to higher SNRs; no change in slope	Shift of psychacoustic curve to higher SNRs; no change in slope
Words in multi-talker babble (**Chapter Three**)	Shift of psychacoustic curve to higher SNRs; no change in slope	Not tested in these cohorts	
Co-ordinate response measure task (**Chapter Four**)	Less spatial release from masking (deficit only at 50º separation in azimuth)		
Spatialised sentences in multi-talker babble (**Chapter Four**)	Overall group deficit; no change in spatial release from masking	Overall group deficit; less spatial release from masking	Overall group deficit; less spatial release from masking
Subjective reports on audio-attentional difficulty (**Chapter Two and Three**)	No impairment	No impairment	Higher subjective scores
Subjective reports on auditory discomfort to non-verbal sounds (**Chapter Two and Three**)	No impairment	No impairment	No impairment
Subjective reports on auditory discomfort to verbal sounds (**Chapter Two and Three**)	No impairment	No impairment	No impairment

EDSS = Expanded disability status scale
Orange cells = Impairment on group level compared to healthy controls
Yellow cells = Tests were not carried out in this cohort
Blue cells = No impairment on group level compared to healthy controls

6.2.4 Deficits in psychoacoustic tasks

MS psychometric discrimination curves which model the relationship between signal-to-noise ratio (SNR) and sentence discrimination accuracy in speech-weighted noise and babble did not change in slope (sentences/dB) but shifted to higher SNRs (dB) compared to controls. This indicates that pwMS required louder target signals to achieve normal discrimination accuracy (**Chapter Two**). Further, the magnitude of the shift in the curve systematically increased with greater disease severity (as measured by EDSS), providing preliminary evidence that SNR may serve as a biomarker for monitoring disease burden (**Figure 1**). Robust longitudinal data in larger cohorts of participants would be the next logical step to confirm disease monitoring efficacy, as the MS disease course develops differently within individuals over time.

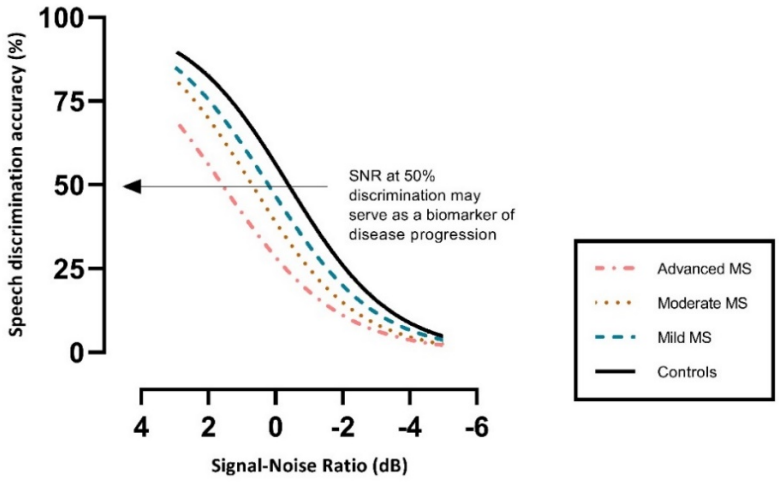


Figure 1: Psychometric discrimination curves of control, mild (Expanded Disability Status Scale (EDSS) score ≤ 1.5), moderate (EDSS score between 2-4.5), and advanced (EDSS score between 5-7) multiple sclerosis (MS) groups. Curves model the relationship between signal-to-noise ratio (dB (decibels)) and speech discrimination accuracy in multi-talker babble (out of 100%). Whilst the curves had similar slopes (sentences/dB), signal-to-noise ratio at 50% discrimination accuracy may provide a valid biomarker of disease progression.

Speech and noise are very rarely localized and more likely to emanate from different directions in realistic acoustic environments, and **Chapter Four** investigated how pwMS use spatial cues to improve SiN intelligibility. In contrast to **Chapters Two and Three**, SiN processing was carried out at a single SNR with spatial separation in the azimuth plane the manipulated variable. In general, MS listeners benefited from spatial cues as group-level discrimination accuracy improved when target speech was separated from the masker by 20º and 50º azimuth compared to 0º (co-localized). However, the mildly affected pwMS group (EDSS ≤ 1.5) received less spatial advantage than healthy counterparts in the Coordinate Response Measure (CRM) task. Specifically, mildly affected pwMS exhibited 10% less spatial release from masking (SRM) than controls when competing speech changed from a colocalized position to 50º separation in azimuth. Error analyses revealed that pwMS were making similar error types to their control counterparts, with the greatest error type recorded the coordinates (colour/number) that belonged to the phrase spoken by the masker talker. This demonstrated that all listeners were able to separate target elements (segregate) but made errors when attributing (or grouping) the coordinates to the correct callsign.

Although we were unable to test all MS disability groups in the CRM, another spatial task utilizing the same multi-talker babble and Bamford-Kowl-Bench sentences in **Chapter Two and Three** was used to evaluate spatial processing in all groups. In the Spatialised Sentences in Noise (SSiN) task, mild, moderate (EDSS 2-4.5) and advanced (EDSS 5-7) MS participants achieved SRM to the same extent as controls, however, all MS groups recalled significantly fewer sentences than controls overall, consistent with the main listener effects described in **Chapter Two and Three**.

In what appears to be a contradictory finding, the CRM sensitively detected an SRM deficit in the mild MS group, whilst the SSiN did not. This is likely to reflect the different parameters used in each test, as the magnitude of SRM depends on a range of factors such as the spatial relationship between target and masker (symmetrical vs asymmetrical configuration) (271), the various parameters of interfering sounds (i.e. number of sounds (258), the fixed SNR etc.) and the similarity between the target and masker (272). Spatial cues become the most salient cue available for parsing sound elements when the acoustic properties of the target speech and masker are high in similarity (251, 267, 273); much like the CRM task which utilizes competing speech with similar syntactic sentence structure, onset/offset timing and same-sex speakers. The SSiN contained more differences between target and masker stimuli, and therefore comprised more readily available segregation cues for the listener to use in addition to spatial cues, perhaps even negating the effect of spatial cues. However, the SSiN task represents a more ecologically-relevant listening environment, much like the 'cocktail party' situation described previously by numerous studies (180, 212, 258). To test our hypothesis, a future investigation could utilize the CRM with systematic changes in parameters: same sex vs. different sex, changes in onset/offset times for competing speech and 1 competing speaker vs. 8 competing speakers. Systematically introducing more readily available segregation cues for the listener will help identify which acoustic cues are the most salient for parsing complex sounds. The presence of several acoustic cues for the listener to use may make spatial cues redundant and if we observe no SRM deficit in conditions where there are more segregation cues readily available, then it provides evidence to support our hypothesis that the CRM is a more sensitive task than the SSiN in detecting spatial processing deficits.

6.2.5 Subjective auditory processing complaints

Despite all MS groups displaying SiN impairments, only the advanced MS group self-reported audio attentional difficulty in daily life (**Chapter Two and Three**). The absence of self-reported auditory difficulty in less severe MS groups could reflect the intrinsic and extrinsic redundancy in auditory processing. The central auditory system itself provides considerable intrinsic redundancy due to multiple pathways and synaptic connections and is thereby likely to be relatively resistant to damage caused by an MS lesion (338). Real-world acoustic environments are also rich with extrinsic redundancy, for example: a listener may turn their head advantageously to increase SNR at one ear and obtain visual cues when environmental sounds are located at different origins. PwMS did not report any auditory discomfort to verbal or non-verbal stimuli on the AADQ; hyperacusis in MS is only described in case reports (223). Refer to Table 1 for a summary of these findings.

6.3 2. The utility of speech-in-noise tests to evaluate sensory and cognitive deficits in MS

The main objective of **Chapter Three** was to evaluate the potential clinical application of our SiN battery in measuring both sensory and cognitive impairment in MS, particularly at the early or mild stages of disease, predating overt physical symptomology. Quantifying neurological deficits in early MS, when little or no physical disability is evident (< EDSS of 3), provides three major benefits: (1) identification of insidious progression that can meaningfully inform prognostication (190, 191), (2) evaluation of treatment effects of potential and current therapeutics very early in the disease, and (3) identification of non-physical problems that have a significant impact on physical, emotional and social aspects of life.

6.3.1 Speech-in-noise paradigms in other populations

SiN paradigms have been routinely used in the elderly, children, and some clinical populations, as it provides a valuable ecological measure of communication disability whilst being quick, easy to administer and non-invasive (339). The premise of testing is that an individual with auditory processing deficits will have difficulty compensating for degraded speech signals as it taxes the auditory system more than speech presented in silence, in contrast, individuals without auditory processing deficits can compensate for degraded signals to some extent.

It is well known that the elderly, over the age of 60, have difficulty in everyday speech communication and function (340). An evaluation of the extensive work done in this field indicates that the effects of age-related decline are likely to be a combination of peripheral, central auditory and higher-processing decline associated with age-related structural changes that normally occur (340). The elderly not only suffer from sensorineural loss, but central auditory processing deficits in binaural processing, temporal resolution, and intensity discrimination, as well as declines in cognitive domains essential for SiN processing, such as attention, working memory and executive processing. Nevertheless, there is considerable variability in the preservation of cognitive reserve during the life span, and language abilities are generally well preserved in older individuals (341). Causal and noncausal interactions between the various factors are complex and intertwined; for example, there is evidence to suggest that the presence of a peripheral hearing loss can accelerate cognitive decline; at the same time, changes in cognition might also be a direct result of using additional resources such as attention and working memory to compensate for sensory loss (333, 334). In other cases, there can be the presence of a central effect of biological ageing without peripheral pathology that also results in a distortion in auditory perception (340). In pwMS, central lesions typically have no effects on pure-tone thresholds, and SiN deficits reported in **Chapters Two, Three and Four** occurred without the presence of peripheral hearing, providing a direct example of the effects of central auditory processing and cognitive deficits on SiN processing in these patients. That being the case, the ageing MS population are also likely to experience peripheral hearing loss, but the additive effects on auditory function are unknown.

In addition to the ageing population, SiN paradigms have been used in clinical populations, such as mild cognitive impairment (MCI) (342, 343) and Autism Spectrum Disorder (ASD) (141). MCI is a condition that involves

a subtle, yet noticeable, decline in cognitive ability that does not meet the criteria for dementia but may precede dementia onset (344). Lee et al., (2016) reported that patients with amnestic mild cognitive impairment (aMCI) did significantly worse than hearing-matched younger and older adults in a SiN task utilizing speech-spectrum noise at low SNRs (high background levels). The investigators also reported that the aMCI group had impairments on the sentence recognition test, not the monosyllabic word test. Reading comprehension and general linguistic competence have been shown to aid sentence in noise intelligibility – not words or syllables which are, in contrast, likely to provide relatively more 'bottom-up' acoustic-phonetic cues (244). Therefore, the authors attributed poorer performance in the aMCI group to impaired lexical semantic processing, which was supported by significant correlations to neuropsychological testing for verbal fluency (342). In our work in pwMS, a clinical population with differing pathology but similarly experiencing cognitive impairment, we found both sentences and words elicited a greater degree of discrimination impairment in MS listeners (**Chapter Three**). A key difference in methodology was that we employed multi-talker babble as the masker for our words-in-noise task, which is likely to induce both informational and energetic masking in comparison to speech-spectrum noise, which is solely energetic masking (refer to Section 6.2.2 on details about the potency of multi-talker babble as a speech masker). Similarly to Lee et al. (2016), we also note significant associations with SiN performances in babble and neuropsychological measures of verbal fluency in MS.

Our lab has previously evaluated the SiN battery used throughout this thesis in normal aged populations (345) and in high-functioning ASD participants (141). ASD is a developmental condition that is characterised by impaired communication skills and repetitive sensory-motor behaviours that can range in severity between individuals (346). Dunlop et al., (2016) provided evidence that high-functioning ASD individuals had poorer performance in typically developing individuals in multi-talker babble, but not in speech-weighted noise. Furthermore, speech-hypersensitivity did not appear to predict performance in the SiN task. Although the results supported the notion that impairments in ASD were likely to reflect an attentional deficit rather a perceptual one, further tests to demonstrate an association between cognitive factors and speech perception were not conducted.

In summary, the use of a SiN test to reflect both sensory and cognitive impairment in pwMS is not an entirely new idea, as it has previously been demonstrated in ageing and clinical populations. However, the body of work in this thesis is the first to systematically evaluate SiN processing in individuals with a wide range of MS severity, and also form associations between neuropsychological factors and SiN performance.

6.3.2 Multi-talker babble: a potent masker for early and mildly affected pwMS

"*It has been said that the best place to hide a leaf is in the forest, and presumably the best place to hide a voice is amongst other voices*" – Miller, 1947

The speech presented in multi-talker babble task sensitively discriminated deficits in early and mild individuals with MS (median EDSS = 0) with normal hearing. In contrast, the speech-weighted noise task did not. Speech-weighted noise is less likely to tax auditory function and reflect the subtle neurological deficits of the mildly impaired group than the more ecologically relevant babble. As Miller suggests: understanding a talker in the presence of other talkers is a particularly difficult listening condition (347, 348).

An important factor in the potency of babble as a masker is the number of talkers. In the studies outlined in **Chapters Two, Three and Four**, babble comprised of eight competing talkers speaking nonsense words, the number confirmed by Simpson and Cooke (2005) to have the most detrimental effect on phoneme detection (232). The general trend described by these authors was that phoneme detection difficulty increased when the number of competing speakers changed from one to eight, but then decreased after eight and up to 512 speakers (232). With such a large number of speakers, babble noise becomes as similar as temporally flat speech-weighted noise, a purely energetic masker of speech sounds (232). Energetic masking refers to competing signals that overlap in frequency and time, creating inaudible effects of target speech (210). The eight-talker babble we used throughout the studies described in this thesis is likely to contain a combination of both energetic and informational masking. Informational masking creates perceptual interference where the competing stimuli are clearly audible, but the speaker is easily confused when attributing sound elements to the target speaker, due to similar-sounding distractors (212, 213). Greater interference occurs with distractors that contain semantic meaning (212) and with native interfering speech (349). **Figure 2** presents a visual analogue of masking properties via images of my pets. **Figure 2A** represents the interference of stimuli that physically overlaps the target image (akin to energetic masking); whilst **Figure 2B** represents the perceptual interference created by similar stimuli (akin to informational masking), where perceptually, the dogs faces blend into one and it becomes more difficult to tease the dog breeds apart. These examples are analogous to competing speech occurring at the same location, i.e. direct overlap of the two dogs. A spatially displaced masker from the target, both in the visual and auditory domain, make more of the target perceivable and thus improves intelligibility, e.g. a shift to the right for **Figure 2C** and **2D** relative to **Figure 2A** and **2B**, respectively.

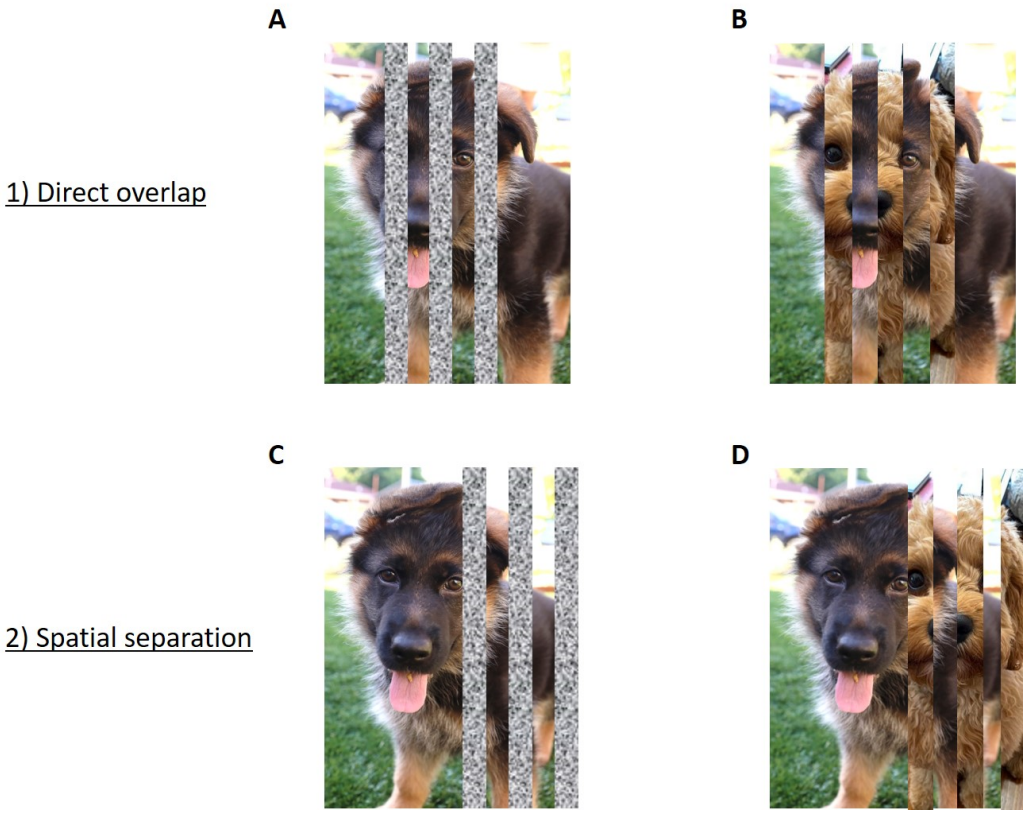


Figure 2: Visual analogue of the masking properties of energetic (A) and informational masking (B), and effects of spatial separation (C&D).

Speech degradation in babble in pwMS might also be related to impaired temporal-resolving capacity. Both Mustillo (1984) and Rappaport (1994) postulated that a contributing factor to SiN deficits in pwMS could be a deficit in temporal processing, possibly related to the delay of signal transmission within the auditory pathways (11, 86). Complex acoustic signals, such as natural masking speech babble, have temporally fluctuating levels, where listeners can use a 'glimpsing' strategy (179) to extract information in "a time-frequency region which contains a reasonably undistorted 'view' of local signal properties" (232). Such fluctuations are not present in the steady state speech-weighted noise, therefore, the temporal resolution skills required to use glimpsing to distinguish speech from babble may partly explain why it is such a potent masker for early MS. This is consistent with the findings of Rappaport et al., (1994) that found that pwMS were significantly impaired in discriminating monosyllables presented monaurally against interrupted noise at every SNR, but not in continuous noise. It was concluded that pwMS, particularly those with forebrain lesions, had impoverished temporal processing deficits, as they could not discriminate speech fragments in the silent periods (86).

Although multi-talker babble is an ecologically relevant sound source often experienced daily by humans, it should be noted that disentangling a complex auditory scene does not always comprise of only speech from humans but also non-living sources such as environmental sounds and human-made machinery. Action sounds produced by 'non-living' sources such as storms, wind, and ocean waves, produce sounds that cannot be fully emulated by a human

listener's motor system; therefore, the human brain relies more on learning and association with visual and tactile inputs to identify and stream auditory objects likely to emanate from the same non-living source. In contrast, sounds produced by living sources activate multisensory and audio-motor networks which convey further intention or meaning behind the sounds (350). Although not studied in the context of this thesis, future investigations into how pwMS process a variety of acoustic scenes comprised of not just speech but other sounds as well, could be useful to better understand and profile the type of central auditory deficits in this population.

6.3.3 Higher order processes required to reconcile perceptual ambiguity

The brain is described as inherently proactive, and not simply reliant on bottom-up sensory information to interpret sensory input (351). When the target signal is degraded because of reduced auditory temporal abilities or attentionally demanding characteristics of babble, MS listeners must place greater demands on cognitive processes to reconcile perceptual ambiguity. Cognitive impairment is now recognized as a prevalent and debilitating component of MS which manifests at all disease stages (64, 70), including onset (189), with the most prominent changes being slowed cognitive processing speed, attention, memory (episodic) and visuospatial skills, and additional impairments in executive function and verbal fluency (57-59). Therefore, it is possible that deficits in SiN processing reflect poorer cognitive control in pwMS. Evidence to support this view can be found in **Chapter Three** where MS listener's susceptibility to speech degradation in babble noise correlated negatively with performance on neuropsychological testing. The remainder of this section will outline the cognitive domains, namely attention and working memory which are particularly pertinent to SiN processing.

A cognitive process of great relevance in SiN processing is attention, in particular, the conscious and active attentional processes that have a robust top-down effect on most of the auditory pathway processes (280). Posner and Peterson (1990) proposed that attentional systems can be divided into three interconnected subsystems: (a) *alerting*, (b) *orientating* and (c) *executive* (281). In the context of SiN processing, *alerting* is driven by bottom-up cues that draw the listener's attention to the salient acoustic cues of the target speech. It requires a state of alertness to prepare attention to an expected signal (282). Participants performing our SiN tasks were given prior instructions to expect target speech to be presented within a constant stream of noise, and perhaps a higher state of alertness would facilitate the detection of speech onset required to focus on the target and improve speech intelligibility.

Orienting refers to the ability to prioritize speech coming from a specific location in space, and reduce the interference caused by a masker at a different location (281, 282), which might be specifically important for our spatial SiN tasks outlined in **Chapter Four**. Spatial cues are particularly useful salient cues for the auditory scene analysis, and several studies have demonstrated that speech recognition in noise improves when the source of the speech is separated horizontally from the interference (249, 251-255).

Executive processes are important for SiN perception and have been previously divided into three subdomains by Dryden et al. (2017) : a) *set-shifting*, b) *inhibitory control* and c) *updating*, or *working memory* (196).

Set-shifting, as the name suggests, is the ability to switch between tasks, or in the context of SiN processing, switch between different target speakers (352). This process may not have been required to perform any of our SiN tasks, however, it could be investigated in pwMS in the future as complex aspects of attention such as selective, divided, and alternating attention are most often impaired in MS – and even CIS - whilst the simplest form, attention span, remains generally intact (59, 61).

Inhibitory control refers to the process of ignoring a distracting interference in order to focus on the desired target (353). The multi-talker babble employed in our SiN tasks comprises of distracting salient features which can involuntarily capture the attention of listeners, and thereby likely to place a greater cognitive load on pwMS who exhibit impaired attention. Poor inhibitory control in pwMS may increase susceptibility to distraction by the informational background masker of babble (196, 231). The CRM task utilizes same-sex speakers and competing phrases with similar syntactic structure and would also heavily engage inhibitory processes to ignore the similar-sounding interfering masker.

Finally, another prominent cognitive process of great relevance in SiN processing is *working memory*, the limited-capacity temporary storage system for active maintenance of information in the face of ongoing processing and distractions (196, 198, 200, 234, 354). In the cognitive hearing sciences, the Ease of Language Understanding (ELU) model (198) emphasizes the subtle balancing act between bottom up and top down aspects of language processing and how and when working memory, in particular, plays a role in the restoration of degraded speech (198, 199). Multifaceted working memory processes integrate lexical and phonological memory stores to 'fill in the gaps' of degraded sounds or mismatches between the perceptual speech input and phonological representations stored in long term memory in order to construct meaningful and coherent speech (198). Working memory impairments are widely reported in early MS (189); some studies have reported that pwMS have deficits associated with maintenance in working memory, whilst others have concluded that the primary deficit is at the level of the central executive which controls and manipulates working memory stores (60). Furthermore, it's worth noting that slowed information processing speed, a particularly prevalent cognitive impairment in MS, is posited to underlie most cognitive changes in MS (236-238). Information processing speed is the speed and efficiency with which information is processed and integrated with other cognitive processes for formation of a behavioural response (236, 237) and has been associated with SiN tasks that utilize long and complex sentences, as rapid comprehensive processing is required for the recruitment of other cognitive processes such as working memory (196).

6.3.4 Neurodegenerative pathology and its association with speech-in-noise performance

In our preliminary study (**Chapter Five**), we explored the association between neurodegenerative processes and SiN performance to determine underlying pathological mechanisms of SiN deficits in pwMS. Brain atrophy, the gradual decline in brain volume attributed to neurodegenerative processes, is approximately 0.5 – 1.35% per year in pwMS, considerably greater than the limits of normal ageing (355). Several mechanisms may underlie neurodegeneration in grey matter (GM), including extensive axonal transection and demyelination leading to retrograde neurodegeneration,

cortical inflammatory pathology, mitochondrial failure, and iron deposition (294, 295). Grey matter pathology has been identified as a significant indicator of cognitive impairment while correlations between cognition and WM pathology have been modest (356).

Previous neuroimaging studies have indicated that more widespread prefrontal and parietal activation outside the primary auditory cortex is associated with SiN compared to clear speech (283, 284). When the auditory cortex in the superior temporal gyrus cannot process speech sounds, higher order functions such as attention, working memory and speech-motoric processes in the prefrontal regions may be recruited to decode impoverished speech (283, 357). However, beyond the temporal lobe, we found no other associations between SiN performance and the volume of other cognitive-associated brain regions as hypothesized. It is plausible that cerebral changes in MS involving adaptive neuroplasticity enable our participants to achieve normal functioning. This is seen as a strong contributing factor to the 'morphological-clinical gap' as morphological insights using MRI only allows assessment of structural changes in the brain, unlike fMRI that detects adaptive changes (298). It is also possible assessing a MS group consisting of a wider spectrum of disability might reveal an association between brain region volumes and SiN processing abilities. Further studies with the incorporation of different neuroimaging techniques and a wider range of disability in patients is required. It should be noted that the nature of this study was highly exploratory and interpretations of correlations between brain structures and SiN performance remain subject to confirmation in future investigations with a larger sample of participants.

6.3.5 Advantages of SiN tasks

SiN assessments have the advantage of being quick and easy to administer and score. As we identified in ***Chapter Three*****, s**entences and words presented in multi-talker babble presented at an intermediate SNR take approximately five minutes to administer and is particularly sensitive to deficits in early and mild disease in MS (**Chapter Three**). SiN testing equipment is also cost effective, minimal, and portable representing minimal resource investment. For these reasons, SiN tasks offer an opportunity for clinicians to screen MS patients for potential cognitive deficits with minimal time and expense, signaling the need for more extensive neuropsychological investigation.

SiN tasks also have high external validity, providing insight into how pwMS navigate their noisy social and work environments in daily life. High external validity also means that individuals are familiar with the task and are therefore more likely to interpret the instructions correctly and complete the task as intended.

6.3.6 Limitations of SiN tasks

It must be acknowledged that our studies are limited to participants with bilateral pure tone averages (PTAs) of hearing threshold levels at 500, 1000, 2000 and 4000 Hz < 25 dB HL. Given the increased likelihood of peripheral hearing loss in older persons and the effects of MS on hearing; the additive concomitant effects of the two may create serious problems understanding SIN in the ageing MS population. Future studies should investigate SiN performance in pwMS with PTAs ≥ 25 dB HL to establish abnormal cutoffs in listeners with varying degrees of hearing loss. If large

data sets are obtained from controls and pwMS with varying degrees of severity, a hearing loss correction could be applied; ensuring that SiN performance in all pwMS can be evaluated and interpreted.

In addition to the exclusion of participants with hearing loss, those with English as their second language (ESL) were also excluded from our study. This limits the use of SiN tasks as a screening tool in all pwMS. Previous work, and our own pilot studies, have demonstrated that ESLs are at a significant disadvantage when trying to discriminate speech from noise. Bilinguals are better able to perceive SiN in their native language due to greater use of higher-level, lingustic context in the native language (358). Developing versions of SiN tasks in other languages may be a way to combat this limitation.

SiN tasks require verbal responses from participants as they must correctly repeat the target sentence heard. This could be particularly problematic if the participant has severe dysarthria, a motor speech disorder that impairs physical production of speech. Intelligibility of the participant's response is an important requirement for the experimenter to correctly score keywords and sentences. Dysarthia affects about 40% of pwMS (359), with the potential for the misinterpretation of SiN deficits as a perceptual problem rather than a production problem. Modifying the task to include a non-verbal response may be an alternative solution for participants with dysarthria.

6.4 Concluding remarks

Acoustic analysis of complex auditory scenes entails both exquisite local neural timing and the integrity of diffuse, higher-level cognitive networks. Assessment of the dynamic interplay between these heterarchical processes sensitively reveals subtle changes in MS that are not reflected in the current gold standard measures in the clinic. Previous studies evaluating auditory function in MS have primarily focussed on using psychacoustic tests that require precise neural timing, using clicks, tones and noise bursts. Unlike these standard laboratory tasks that present no external validity, evaluating speech discrimination in noise not only requires complex temporal resolution but also presents real-life cognitive demands. We propose that SiN tasks could be employed as a screening tool for changes in cognitive function, for monitoring disease activity and possibly evaluating the efficacy of therapeutic strategies. Further refinement to SiN parameters may be achieved in future studies to increase capacity to detect subtle central auditory deficits in pwMS; SiN methodology has the advantage of studying many more systemic variations in SNR, sentence difficulty and saliency of background maskers. We do not suggest that SiN tasks replace more conventional neuropsychological measures used to assess cognitive function in MS or that they might be useful in establishing diagnosis, however, we suggest that they might be employed as a complementary test due to their ease of use, speed and cost effectiveness – features which are advantageous in a clinical setting.

Supplementary Material

They're drinking tea.

Mother opens the drawer.

An old woman was at home.

He dropped his money.

The kitchen window was clean.

The girl plays with the baby.

She's helping her friend.

The children washed the plates.

The postman comes early.

The sign showed the way.

The grass is getting long.

The match fell on the floor.

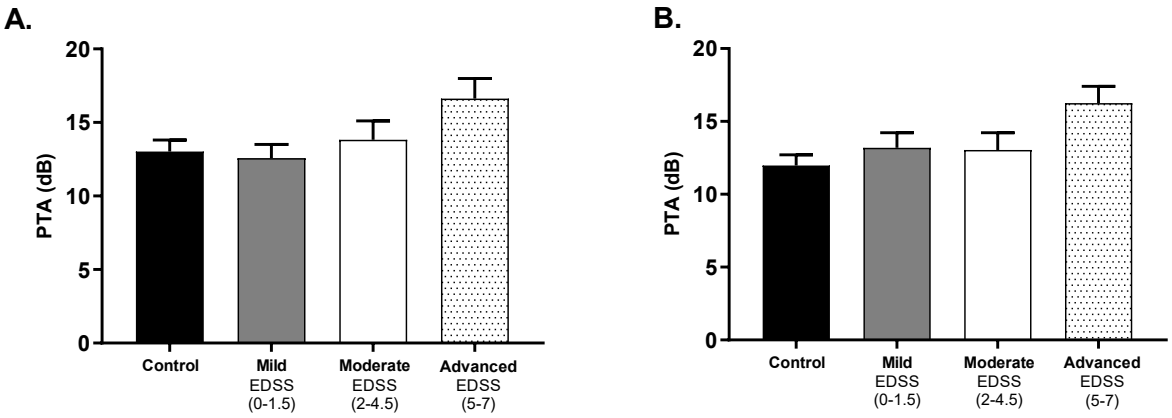
Supplementary Figure A.1. An example of a list of sentences (Bamford-Kowal-Bench Standard List No. 14, copyright Academic Press Ltd). Sentence length did not exceed seven syllables. Key words, which are used for scoring purposes, are underlined. All three keywords had to be correct for a sentence to be scored as correct.

Supplementary Table A.1. All the predictor variables used in the generalised linear mixed effects models (with a logit link function) used to predict correct sentence recall on any trial

Categorical/ordinal variables	Number of unique values	Coded/Range
MS disability group (categorical)	4	Controls = 0; Mild = 1; Moderate = 2; Advanced = 3
Masker Type (categorical)	2	SWN = 0; BN" = 1
Sex (categorical)	2	Female = 0; Male = 1
Disease phenotype (categorical)	2	RR = 0; SP = 1
EDSS score (ordinal)	14	0 – 7 (0.5 intervals)
Signal to noise ratio (ordinal)	6	-7, -5, -3, -1 & 1 in SWN; -5, -3, -1, 1, 3 & 7
Trial order (ordinal)	10	1-10 for each SNR

Demographic variables	Mean	SD	Range
Age (years)	47.14	10.27	24 – 64
Education (years)	14.08	2.18	11 - 20
Disease Duration (years)	13.80	7.79	1 - 32
Four pure tone threshold average (dB HL")	L ear = 13.02 R ear = 13.75	L ear = 4.63 R ear = 4.97	L ear = 2.5 – 18.75 R ear = 2.5 – 22.5

Note: Random variable = random intercepts for between-participant heterogeneity
Outcome variable =sentence incorrect = 0; sentence correct = 1 (binomial distribution)
SD = standard deviation
SWN = speech-weighted noise
BN = Babble noise
RR = Relapsing Remitting
SP = Secondary Progressive
EDSS = Expanded Disability Status Scale
dB HL = decibels Hearing Level
L = left
R = right



Supplementary Figure A.2. MS disability groups had similar hearing sensitivity to controls. The mean pure tone averages (PTA) ± standard error of the mean (SEM) obtained for left (A) and right (B) ears in controls (black; n=38) and pwMS with mild (grey; n=20), moderate (white; n=16) and advanced (patterned; n=10) disability. PTAs were calculated as the average of thresholds (dB HL) at 500, 1000, 2000 and 4000Hz. No statistical differences in PTAs were found between groups (One-way ANOVA).

Supplementary Table A.2. Goodness of fit (R^2), slope and midpoint values (± SEM) for Boltzmann sigmoidal functions for the performance of controls and MS with mild, moderate, and advanced impairment in the sentences in speech-weighted task.

	R^2	Midpoint ± SE (dB)	Slope ± SE (sentence/dB)
Controls	0.95	-6.79 ± 0.19	1.54 ± 0.13
MS; Mild disability (EDSS 0-1.5; Median = 0)	0.93	-6.38 ± 0.22	1.55 ± 0.15
MS; Moderate disability (EDSS 2-4.5; Median = 2.5)	0.92	-5.86 ± 0.31*	1.96 ± 0.14
MS; Advanced disability (EDSS 5-7; Median = 6)	0.89	-4.85 ± 0.43****	1.92 ± 0.2

* ($p < 0.05$); **** ($p < 0.0001$) compared to controls (One Way ANOVA, Tukey's post hoc test)
The top and bottom of the Boltzmann sigmoidal functions were constrained to 10 and 0 respectively

Supplementary Table A.3. Goodness of fit (R^2), slope and midpoint values (± SEM) for Boltzmann sigmoidal functions for the performance of controls and MS with mild, moderate and advanced impairment in the sentences in babble task.

	R^2	Midpoint ± SE (dB)	Slope ± SE (sentence/dB)
Controls	0.95	-0.39 ± 0.38	1.43 ± 0.08
MS; Mild disability (EDSS 0-1.5; Median = 0)	0.94	0.27 ± 0.2*	1.57 ± 0.11
MS; Moderate disability (EDSS 2-4.5; Median = 2.5)	0.92	0.75 ± 0.2**	1.46 ± 0.13
MS; Advanced disability (EDSS 5-7; Median = 6)	0.94	1.45 ± 0.31***	1.42 ± 0.25

* ($p < 0.05$); ** ($p < 0.001$) **** ($p < 0.0001$) compared to controls (One Way ANOVA, Tukey's post hoc test)
The top and bottom of the Boltzmann sigmoidal functions were constrained to 10 and 0 respectively

Auditory Attention and Difficulty Questionnaire (AADQ)

Item number and statement

1. I find the sound of doorbells annoyingly loud
2. I find the sound of a telephone ringing to be uncomfortably loud
3. I find restaurants and cafes to be uncomfortably loud.
4. I have arguments with my family or friends because I think they talk too loudly.
5. I have the TV or radio volume much lower than do my family or friends
6. I find supermarkets to be uncomfortably loud
7. The sounds of running water, like a toilet or shower, are uncomfortably loud
8. I can hear the sounds of birds singing in the mornings
9. The sounds of building work are painfully loud
10. Traffic noises are uncomfortably loud
11. The sound of screeching tyres is uncomfortably loud.
12. When I am in a theatre watching a movie or play, I find it uncomfortably loud when people around me are whispering and rustling packets
13. I have trouble understanding others when an air conditioner or fan is on
14. Unexpected sounds, like a smoke detector or alarm bell, are uncomfortable
15. I avoid social gatherings (like parties) because I find the noise levels annoying
16. I find parties are too loud to be able to concentrate to have a conversation
17. I can understand conversations even when several people are talking
18. I have difficulty hearing a conversation when I'm with one of my family at home
19. I have difficulty following a conversation on the phone or mobile when I'm at home
20. When I am having a quiet conversation with a friend, I have difficulty understanding them
21. When I'm seeing my doctor in his/her rooms, it is hard to follow the conversation
22. I have trouble understanding dialogue in a movie or at the theatre
23. When I am talking with someone across a large empty room, I have difficulty understanding what they say
24. I miss a lot of information when I'm listening to a lecture or a public talk
25. When a speaker is addressing a small group, and everyone is listening quietly, I have to strain to hear
26. I have to ask people to repeat themselves in one-on-one conversations in a quiet room
27. I have trouble understanding a waiter/waitress in a quiet restaurant
28. When I am in a small office, talking or answering questions, I have difficulty following the conversation
29. When I am having dinner with several other people, I have difficulty following the conversation because I find it hard to identify who is speaking
30. I have difficulty communicating with others when we are in a crowd
31. When I am in a crowded supermarket talking with the cashier, I can follow the conversation*
32. I have difficulty understanding a shop assistant in a crowded shop

33. In social situations I often feel left out because people think I have difficulty following the conversations

Supplementary Figure A.3. The 33 items on the Auditory Attention and Difficulty Questionnaire (AADQ) were Modified from the University of Auckland Evaluation of Hearing Performance, the Amsterdam Inventory, The Denver Scales (Schow & Nerbonne JARA 1980), and the Hearing Handicap Inventory for Adults – Screening (Ventry, I. & Weinstein, B. Ear Hear 1982). Statements were summarized into three components: Component 1, the Audio-Attentional Difficulty subscale, measured difficulties attending to speech in noisy environments from fourteen items (items 18-30, 33). Component 2, the Auditory Discomfort (Non-Verbal) subscale, measured discomfort to non-verbal environmental sounds from eight items (items 1, 2, 4, 9-12, 14). Component 3, the Auditory Discomfort (Verbal) subscale, measured discomfort to verbal sounds from seven items (items 3, 6, 13, 15, 16, 31, 32). The questionnaire was completed on paper during the testing session under no time restriction.
* Item 31 had negative valence and so participants' responses were reversed

Supplementary Table A.4. Participant characteristics

		Control	Mild MS	Moderate MS	Advanced MS	*p*
	Number of participants	20	23	16	8	
Demographics	**Sex F(M)**	17(3)	24(3)	13(3)	7(1)	
	Age, (yrs)					
	Mean (SD)	45.85(10.82)	47.03(8.73)	46.67(11.90)	48.13(6.22)^	0.98 [a]
	Range	28 - 60	28 - 65	28 – 64	36 – 58	
Audiometry	**Pure tone average (dB HL)**					
	Left (Mean, SD)	13.12(5.90)	12.46(4.67)	13.83(5.11)	16.88(4.77)	0.22 [b]
	Range	-1.25 to 18.75	2.5 to 18.8	6.25 to 18.75	10 to 23.75	
	Right (Mean, SD)	12.06(5.13)	13.59(5.37)	13.05(4.70)	16.09(3.81)	0.29 [b]
	Range	2.5 to 23.75	2.5 to 22.5	5 to 22.5	10 to 20	
Disease Characteristics	**Disease duration (yrs)**	-				
	Mean (SD)	-	10.45(5.69)	13.85(8.46)	17.88(7.90)	**0.03** [b]
	Range	-	1 - 22	1 - 32	6 – 31	
	***EDSS**	-				
	Mean(SD)	-	0.11(0.37)	2.30(1.38)	6.10(0.62)	**< 0.0001** [a]
	Range	-	0 - 1.5	2 – 4.5	5 - 7	
	Phenotype RR(SP)		23(0)	13(3)	2(6)	
	On disease modifying therapy (n, %)	NA	91.3%	93.7%	75%	0.34 [c]

*EDSS = Expanded Disability Status Scale Score determined by a neurologist within 6 months of audiological testing;
F = female; M = male
SD = Standard deviation
dB HL = decibels hearing level
RR = Relapsing Remitting
SP = Secondary Progressive

[a] Kruskal-Wallis Test
[b] One-Way ANOVA
[c] Chi-squared Test
^ missing data from one participant
NOTE: Demographics and audiometry were compared between controls, mild, moderate and advanced MS. Disease characteristics were compared between mild, moderate and advanced MS

References

1. Feigin VL, Abajobir AA, Abate KH, Abd-Allah F, Abdulle AM, Abera SF, et al. Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet Neurology. 2017;16(11):877-97.
2. Leray E, Moreau T, Fromont A, Edan G. Epidemiology of multiple sclerosis. Revue Neurologique. 2016;172(1):3-13.
3. Leray E, Yaouanq J, Le Page E, Coustans M, Laplaud D, Oger J, et al. Evidence for a two-stage disability progression in multiple sclerosis. Brain. 2010;133(7):1900-13.
4. Lublin FDMD, Reingold SCP, National Multiple Sclerosis Society Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Defining the clinical course of multiple sclerosis: Results of an international survey. Neurology. 1996;46(4):907-11.
5. Lucchinetti C, Brück W, Parisi J, Scheithauer B, Rodriguez M, Lassmann H. Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. Annals of Neurology. 2000;47(6):707-17.
6. Ebers GC. Natural history of multiple sclerosis. Journal of Neurology, Neurosurgery and Psychiatry. 2001;71(6):ii16.
7. Fielding J, Clough M. Degenerative disorders of the brain: London : Routledge, Taylor & Francis Group; 2019. 163-86 p.
8. Cosh A, Carslaw H. Multiple sclerosis: symptoms and diagnosis. InnovAiT. 2014;7(11):651-7.
9. Stach BA, Delgado-Vilches G, Smith-Farach S. Hearing Loss in Multiple Sclerosis. Semin Hear. 1990;11(03):221-30.
10. Furst M, Levine RA. Chapter 36 - Hearing disorders in multiple sclerosis. In: Michael J. Aminoff FB, Dick FS, editors. Handbook of Clinical Neurology. Volume 129: Elsevier; 2015. p. 649-65.
11. Mustillo P. Auditory Deficits in Multiple Sclerosis: A Review. Audiology. 1984;23(2):145-64.
12. Lewis MS, Wilmington D, Hutter M, McMillan G, Casiana L, Fitzpatrick M, et al. Preliminary Identification of Central Auditory Processing Screening Tests for Individuals with Multiple Sclerosis. Semin Hear. 2012;33(03):261-73.
13. Arya A. Multiple Sclerosis: A review. International Journal of Pharmaceutical Sciences and Research. 2010;2(1):13-23.
14. Goldenberg MM. Multiple Sclerosis Review. Pharmacy and Therapeutics. 2012;37(3):175-84.
15. McDonald WI. The dynamics of multiple sclerosis. Journal of Neurology. 1993;240(1):28-36.
16. Charcot JM. Histology de la sclerose en plaques. Gaz Hop. 1868;41:554-5, 7-8, 66.
17. Gelfand JM. Chapter 12 - Multiple sclerosis: diagnosis, differential diagnosis, and clinical presentation. In: Douglas SG, editor. Handbook of Clinical Neurology. Volume 122: Elsevier; 2014. p. 269-90.
18. Weissert R. The Immune Pathogenesis of Multiple Sclerosis. Journal of Neuroimmune Pharmacology. 2013;8(4):857-66.

19. Lassmann H. Pathophysiology of inflammation and tissue injury in multiple sclerosis: What are the targets for therapy. Journal of the Neurological Sciences. 2011;306(1):167-9.
20. Lassmann H, Brück W, Lucchinetti CF. The Immunopathology of Multiple Sclerosis: An Overview. Brain Pathology. 2007;17(2):210-8.
21. Hohlfeld R, Wekerle H. Autoimmune concepts of multiple sclerosis as a basis for selective immunotherapy: From pipe dreams to (therapeutic) pipelines. Proceedings of the National Academy of Sciences. 2004;101(suppl 2):14599-606.
22. Navikas V, Link H. Review: cytokines and the pathogenesis of multiple sclerosis. Journal of Neuroscience Research. 1996;45(4):322-33.
23. von Büdingen HC, Bar-Or A, Zamvil SS. B cells in multiple sclerosis: connecting the dots. Current Opinion in Immunology. 2011;23(6):713-20.
24. Wekerle H. B cells in multiple sclerosis. Autoimmunity. 2017;50(1):57-60.
25. Bernard CCA, de Rosbo NK. Multiple sclerosis: an autoimmune disease of multifactorial etiology. Current Opinion in Immunology. 1992;4(6):760-5.
26. Rejdak K, Jackson S, Giovannoni G. Multiple sclerosis: a practical overview for clinicians. British Medical Bulletin. 2010;95(1):79-104.
27. Milo R, Kahana E. Multiple sclerosis: Geoepidemiology, genetics and the environment. Autoimmunity Reviews. 2010;9(5):A387-A94.
28. Zajicek J. The epidemiology of multiple sclerosis. Journal of Neurology. 2007;254(12):1742-.
29. McDonald WI, Compston A, Edan G, Goodkin D, Hartung H-P, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis. Annals of Neurology. 2001;50(1):121-7.
30. Polman CH, Reingold SC, Edan G, Filippi M, Hartung H-P, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". Annals of Neurology. 2005;58(6):840-6.
31. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. Annals of Neurology. 2011;69(2):292-302.
32. Matsushita T, Tateishi T, Isobe N, Yonekawa T, Yamasaki R, Murai H, et al. Characteristic Cerebrospinal Fluid Cytokine/Chemokine Profiles in Neuromyelitis Optica, Relapsing Remitting or Primary Progressive Multiple Sclerosis. PLoS One. 2013;8(4):e61835.
33. Wolinsky JS. The diagnosis of primary progressive multiple sclerosis. Journal of the neurological sciences. 2003;206(2):145-52.
34. Hurwitz BJ. The diagnosis of multiple sclerosis and the clinical subtypes. Annals of Indian Academy of Neurology. 2009;12(4):226-30.
35. Miller DH, Chard DT, Ciccarelli O. Clinically isolated syndromes. The Lancet Neurology. 2012;11(2):157-69.
36. Tintore M, Rovira A, Rio J, Nos C, Grive E, Tellez N, et al. Baseline MRI predicts future attacks and disability in clinically isolated syndromes. Neurology. 2006;67(6):968-72.

37. Beck RW, Trobe JD, Moke PS, Gal RL, Xing D, Bhatti MT, et al. High- and low-risk profiles for the development of multiple sclerosis within 10 years after optic neuritis: experience of the optic neuritis treatment trial. Archives of ophthalmology (Chicago, Ill : 1960). 2003;121(7):944-9.
38. Tintore M, Rovira A, Rio J, Tur C, Pelayo R, Nos C, et al. Do oligoclonal bands add information to MRI in first attacks of multiple sclerosis? Neurology. 2008;70(13 Pt 2):1079-83.
39. Avasarala JR. Clinically isolated syndrome — Rethinking the diagnosis. Journal of the Neurological Sciences. 2015;353(1–2):79-80.
40. Lublin FD, Reingold SC, Cohen JA, Cutter GR, Sørensen PS, Thompson AJ, et al. Defining the clinical course of multiple sclerosis: The 2013 revisions. Neurology. 2014;83(3):278-86.
41. Fambiatos A, Jokubaitis V, Horakova D, Kubala Havrdova E, Trojano M, Prat A, et al. Risk of secondary progressive multiple sclerosis: A longitudinal study. Multiple sclerosis. 2020;26(1):79-90.
42. O'Loughlin E, Hourihan S, Chataway J, Playford ED, Riazi A. The experience of transitioning from relapsing remitting to secondary progressive multiple sclerosis: views of patients and health professionals. Disability and rehabilitation. 2017;39(18):1821-8.
43. Lublin FD. New multiple sclerosis phenotypic classification. European neurology. 2014;72 Suppl 1:1-5.
44. Sá MJ. Physiopathology of symptoms and signs in multiple sclerosis. Arquivos de Neuro-Psiquiatria. 2012;70:733-40.
45. Compston A, Coles A. Multiple sclerosis. The Lancet.372(9648):1502-17.
46. Balcer LJ, Miller DH, Reingold SC, Cohen JA. Vision and vision-related outcome measures in multiple sclerosis. Brain. 2015;138(Pt 1):11-27.
47. Costello F. Vision Disturbances in Multiple Sclerosis. Semin Neurol. 2016;36(02):185-95.
48. Motl RW, Learmonth YC. Neurological disability and its association with walking impairment in multiple sclerosis: brief review. Neurodegenerative Disease Management. 2014;4:491+.
49. Balcer LJ. Optic Neuritis. New England Journal of Medicine. 2006;354(12):1273-80.
50. Pedro Barros MJS. Management of Motor Symptoms in Multiple Sclerosis. US Neurology. 2013;9(1):24-9.
51. Kalron A. Gait variability across the disability spectrum in people with multiple sclerosis. Journal of the Neurological Sciences. 2016;361:1-6.
52. Rivera VM. The Nature of Multiple Sclerosis. Semin Hear. 1990;11(03):207-19.
53. Rao SML, G J ; Bernardin, L ; Unverzagt, F. Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. Neurology. 1991;41(5):685-91.
54. Rao SM, Leo GJ, Ellington L, Nauertz T, Bernardin L, Unverzagt F. Cognitive dysfunction in multiple sclerosis.: II. Impact on employment and social functioning. Neurology. 1991;41(5):692-6.
55. Rao SM. Neuropsychology of multiple sclerosis: a critical review. J Clin Exp Neuropsychol. 1986;8(5):503-42.
56. Calabrese M, Rinaldi F, Grossi P, Gallo P. Cortical pathology and cognitive impairment in multiple sclerosis. Expert Review of Neurotherapeutics. 2011;11(3):425-32.

57. Sumowski JF, Benedict R, Enzinger C, Filippi M, Geurts JJ, Hamalainen P, et al. Cognition in multiple sclerosis: State of the field and priorities for the future. Neurology. 2018;90(6):278-88.
58. Oreja-Guevara C, Ayuso Blanco T, Brieva Ruiz L, Hernandez Perez MA, Meca-Lallana V, Ramio-Torrenta L. Cognitive Dysfunctions and Assessments in Multiple Sclerosis. Front Neurol. 2019;10(581):581.
59. Amato MP, Portaccio E, Goretti B, Zipoli V, Hakiki B, Giannini M, et al. Cognitive impairment in early stages of multiple sclerosis. Neurol Sci. 2010;31(Suppl 2):S211-4.
60. DeLuca J, Chelune GJ, Tulsky DS, Lengenfelder J, Chiaravalloti ND. Is Speed of Processing or Working Memory the Primary Information Processing Deficit in Multiple Sclerosis? Journal of Clinical and Experimental Neuropsychology. 2004;26(4):550-62.
61. Amato MP, Zipoli V, Portaccio E. Cognitive changes in multiple sclerosis. Expert Rev Neurother. 2008;8(10):1585-96.
62. DeSousa EA, Albert RH, Kalman B. Cognitive impairments in multiple sclerosis: a review. American journal of Alzheimer's disease and other dementias. 2002;17(1):23-9.
63. Oset M, Stasiolek M, Matysiak M. Cognitive Dysfunction in the Early Stages of Multiple Sclerosis-How Much and How Important? Current neurology and neuroscience reports. 2020;20(7):22-.
64. Ruet AMD, Deloire MP, Charre-Morin JM, Hamel DM, Brochet BMD. Cognitive impairment differs between primary progressive and relapsing-remitting MS. Neurology. 2013;80(16):1501-8.
65. Zakzanis KK. Distinct Neurocognitive Profiles in Multiple Sclerosis Subtypes. Archives of Clinical Neuropsychology. 2000;15(2):115-36.
66. Staffen W, Mair A, Zauner H, Unterrainer J, Niederhofer H, Kutzelnigg A, et al. Cognitive function and fMRI in patients with multiple sclerosis: evidence for compensatory cortical activation during an attention task. Brain. 2002;125(6):1275-82.
67. Audoin B, Duong MVA, Ranjeva JP, Ibarrola D, Malikova I, Confort-Gouny S, et al. Magnetic resonance study of the influence of tissue damage and cortical reorganization on PASAT performance at the earliest stage of multiple sclerosis. Human Brain Mapping. 2005;24(3):216-28.
68. Forn C, Barros-Loscertales A, Escudero J, Belloch V, Campos S, Parcet MA, et al. Cortical reorganization during PASAT task in MS patients with preserved working memory functions. NeuroImage. 2006;31(2):686-91.
69. Denney DR, Sworowski LA, Lynch SG. Cognitive impairment in three subtypes of multiple sclerosis. Archives of Clinical Neuropsychology. 2005;20(8):967-81.
70. Brissart H, Morele E, Baumann C, Perf ML, Leininger M, Taillemite L, et al. Cognitive impairment among different clinical courses of multiple sclerosis. Neurol Res. 2013;35(8):867-72.
71. Huang WJ, Chen WW, Zhang X. Multiple sclerosis: Pathology, diagnosis and treatments. Exp Ther Med. 2017;13(6):3163-6.
72. Moore DR. Auditory processing disorder (APD): Definition, diagnosis, neural basis, and intervention. Audiological Medicine. 2006;4(1):4-11.
73. Celesia GG. Chapter 29 - Hearing disorders in brainstem lesions. In: Michael J. Aminoff FB, Dick FS, editors. Handbook of Clinical Neurology. Volume 129: Elsevier; 2015. p. 509-36.

74. Hackett TA. Chapter 2 - Anatomic organization of the auditory cortex. In: Michael J. Aminoff FB, Dick FS, editors. Handbook of Clinical Neurology. Volume 129: Elsevier; 2015. p. 27-53.
75. Pickles JO. Chapter 1 - Auditory pathways: anatomy and physiology. In: Michael J. Aminoff FB, Dick FS, editors. Handbook of Clinical Neurology. Volume 129: Elsevier; 2015. p. 3-25.
76. Hellmann MA, Steiner I, Mosberg-Galili R. Sudden sensorineural hearing loss in multiple sclerosis: clinical course and possible pathogenesis. Acta Neurologica Scandinavica. 2011;124(4):245-9.
77. Moncrieff D, Jorgensen L, Ortmann A. Chapter 11 - Psychophysical auditory tests. In: Gastone GC, editor. Handbook of Clinical Neurophysiology. Volume 10: Elsevier; 2013. p. 217-34.
78. Lewis MS, Lilly DJ, Hutter MM, Bourdette DN, McMillan GP, Fitzpatrick MA, et al. Audiometric hearing status of individuals with and without multiple sclerosis. Journal of Rehabilitation Research & Development. 2010;47(7):669-78.
79. Noffsinger D. Auditory and Vestibular Aberrations in Multiple Sclerosis: Almqvist & Wiksell; 1972.
80. Daugherty WT, Lederman RJ, Nodar RH, Conomy JP. HEaring loss in multiple sclerosis. Archives of Neurology. 1983;40(1):33-5.
81. Doty RL, Tourbier I, Davis S, Rotz J, Cuzzocreo JL, Treem J, et al. Pure-Tone Auditory Thresholds Are Not Chronically Elevated in Multiple Sclerosis. Behavioral Neuroscience. 2012;126(2):314-24.
82. Yamasoba T, Sakai K, Sakurai M. Role of acute cochlear neuritis in sudden hearing loss in multiple sclerosis. Journal of the Neurological Sciences. 1997;146(2):179-81.
83. Fischer C, Mauguière F, Ibanez V, Confavreux C, Chazot G. The acute deafness of definite multiple sclerosis: Baep patterns. Electroencephalography and Clinical Neurophysiology. 1985;61(1):7-15.
84. Atula S, Sinkkonen ST, Saat R, Sairanen T, Atula T. Association of multiple sclerosis and sudden sensorineural hearing loss. Multiple Sclerosis Journal – Experimental, Translational and Clinical. 2016;2:2055217316652155.
85. Leite HFB, Leite JdCB, Melo MHdA, Vasconcelos CCF, Alvarenga RMdP. Deafness in Patients with Multiple Sclerosis. Audiology and Neurotology. 2014;19(4):261-6.
86. Rappaport JM, Gulliver JM, Phillips DP, Van Dorpe RA, Maxner CE, Bhan V. Auditory temporal resolution in multiple sclerosis. The Journal of otolaryngology. 1994;23(5):307-24.
87. de Seze J, Assouad R, Stojkovic T, Desaulty A, Dubus B, Vermersch P. [Hearing loss in multiple sclerosis: clinical, electrophysiologic and radiological study]. Rev Neurol (Paris). 2001;157(11 Pt 1):1403-9.
88. Cure JK, Cromwell LD, Case JL, Johnson GD, Musiek FE. Auditory dysfunction caused by multiple sclerosis: detection with MR imaging. AJNR Am J Neuroradiol. 1990;11(4):817-20.
89. Barratt HJ, Miller, D., & Rudge, P. The site of lesion causing deafness in multiple sclerosis. Scandinavian Audiology. 1988;17:67-71.
90. Tabira T, Tsuji S, Nagashima T, Nakajima T, Kuroiwa Y. Cortical deafness in multiple sclerosis. Journal of Neurology, Neurosurgery & Psychiatry. 1981;44(5):433.

91. Shea JJ, 3rd, Brackmann DE. Multiple sclerosis manifesting as sudden hearing loss. Otolaryngology--head and neck surgery : official journal of American Academy of Otolaryngology-Head and Neck Surgery. 1987;97(3):335-8.

92. Helzner EP, Cauley JA, Pratt SR, Wisniewski SR, Zmuda JM, Talbott EO, et al. Race and Sex Differences in Age-Related Hearing Loss: The Health, Aging and Body Composition Study. Journal of the American Geriatrics Society. 2005;53(12):2119-27.

93. Seidman MD, Ahmad N, Joshi D, Seidman J, Thawani S, Quirk WS. Age-related hearing loss and its association with reactive oxygen species and mitochondrial DNA damage. Acta oto-laryngologica Supplementum. 2004(552):16-24.

94. Sharashenidze N, Schacht J, Kevanishvili Z. Age-related hearing loss: gender differences. Georgian medical news. 2007(144):14-8.

95. Simpkins WT, Jr. An audiometric profile in multiple sclerosis. Archives of Otolaryngology. 1961;73(5):557-64.

96. Wilson RH, Noe CM, Cruickshanks KJ, Wiley TL, Nondahl DM. Prevalence and degree of hearing loss among males in Beaver Dam cohort: comparison of veterans and nonveterans. Journal of Rehabilitation Research & Development. 2010;47:505+.

97. Musiek FE, G, Gollegly, K. M., Kibbe, K.S. and Reeves, A.G. Electrophysiologic and Behavioral Auditory Findings in Multiple Sclerosis. The American Journal of Otology. 1989;10(5):343-50.

98. (AAA) AAoA. Diagnosis, treatment and management of children and adults with central auditory processing disorder. 2010.

99. Audiology BSo. Position Statement and Practice Guidance; Auditory Processing Disorder (APD). 2017.

100. Weihing J, Chermak G, Musiek F. Auditory Training for Central Auditory Processing Disorder. 2015;36(4):199-215.

101. Nagao K, Riegner T, Padilla J, Greenwood LA, Loson J, Zavala S, et al. Prevalence of Auditory Processing Disorder in School-Aged Children in the Mid-Atlantic Region. Journal of the American Academy of Audiology. 2016;27(9):691-700.

102. Ocak E, Eshraghi RS, Danesh A, Mittal R, Eshraghi AA. Central Auditory Processing Disorders in Individuals with Autism Spectrum Disorders. Balkan Med J. 2018;35(5):367-72.

103. Legatt AD. Chapter 17 - Electrophysiologic auditory tests. In: Michael J. Aminoff FB, Dick FS, editors. Handbook of Clinical Neurology. Volume 129: Elsevier; 2015. p. 289-311.

104. Robinson K, Rudge P. Auditory evoked responses in multiple sclerosis. Lancet. 1975;1(7917):1164-6.

105. Matathias O, Sohmer H, Biton V. Central Auditory Tests and Auditory Nerve-Brainstem Evoked Responses in Multiple Sclerosis. Acta Oto-Laryngologica. 1985;99(3-4):369-76.

106. Matas CG, Matas SL, Oliveira CR, Goncalves IC. Auditory evoked potentials and multiple sclerosis. Arq Neuropsiquiatr. 2010;68(4):528-34.

107. Çelebisoy N, Aydo I, Ekmekçi Ö, Akürekli Ö. Middle latency auditory evoked potentials (MLAEPs) in (MS). Acta Neurologica Scandinavica. 1996;93(5):318-21.

108. Ivica N, Titlic M, Pavelin S. P300 Wave Changes in Patients with Multiple Sclerosis. Acta Informatica Medica. 2013;21(3):205-7.
109. Magnano I, Aiello I, Piras MR. Cognitive impairment and neurophysiological correlates in MS. Journal of the Neurological Sciences. 2006;245(1):117-22.
110. Chiappa KH, Harrison JL, Brooks EB, Young RR. Brainstem auditory evoked responses in 200 patients with multiple sclerosis. Annals of Neurology. 1980;7(2):135-43.
111. Invernizzi P, Bertolasi L, Bianchi M, Turatti M, Gajofatto A, Benedetti M. Prognostic value of multimodal evoked potentials in multiple sclerosis: the EP score. Official Journal of the European Neurological Society. 2011;258(11):1933-9.
112. Robinson K, Rudge P. Abnormalities of the auditory evoked potentials in patients with multiple sclerosis. Brain. 1977;100 Pt 1:19-40.
113. Kofler B, Oberascher G, Pommer B. Brain-stem involvement in multiple sclerosis: a comparison between brain-stem auditory evoked potentials and the acoustic stapedius reflex. Journal of Neurology. 1984;231(3):145-7.
114. Japaridze G, Shakarishvili R, Kevanishvili Z. Auditory brainstem, middle-latency, and slow cortical responses in multiple sclerosis. Acta Neurologica Scandinavica. 2002;106(1):47-53.
115. Goldstein R, Rodman LB. Early Components of Averaged Evoked Responses to Rapidly Repeated Auditory Stimuli. Journal of Speech, Language, and Hearing Research. 1967;10(4):697-705.
116. Jerger JF, Oliver TA, Chmiel RA, Rivera VM. Patterns of Auditory Abnormality in Multiple Sclerosis. Semin Hear. 1990;11(03):276-90.
117. Sutton S, Braren M, Zubin J, John ER. Evoked-Potential Correlates of Stimulus Uncertainty. Science. 1965;150(3700):1187-8.
118. Avan P, Giraudet F, Buki B. Importance of binaural hearing. Audiology & neuro-otology. 2015;20 Suppl 1:3-6.
119. Tollin DJ, Yin TCT. Sound Localization: Neural Mechanisms. In: Squire LR, editor. Encyclopedia of Neuroscience. Oxford: Academic Press; 2009. p. 137-44.
120. Furst M, Levine RA, McGaffigan PM. Click lateralization is related to the β component of the dichotic brainstem auditory evoked potentials of human subjects. The Journal of the Acoustical Society of America. 1985;78(5):1644-51.
121. Furst M, Eyal S, Korczyn AD. Prediction of binaural click lateralization by brainstem auditory evoked potentials. Hearing Research. 1990;49(1–3):347-59.
122. Levine RA, Gardner JC, Fullerton BC, Stufflebeam SM, Carlisle EW, Furst M, et al. Effects of multiple sclerosis brainstem lesions on sound lateralization and brainstem auditory evoked potentials. Hearing Research. 1993;68(1):73-88.
123. Furst M, Levine RA, Korczyn AD, Fullerton BC, Tadmor R, Algom D. Brainstem lesions and click lateralization in patients with multiple sclerosis. Hearing Research. 1995;82(1):109-24.
124. Aharonson V, Furst M, Levine RA, Chaigrecht M, Korczyn AD. Lateralization and binaural discrimination of patients with pontine lesions. The Journal of the Acoustical Society of America. 1998;103(5):2624-33.

125. Furst M, Aharonson V, Levine RA, Fullerton BC, Tadmor R, Pratt H, et al. Sound lateralization and interaural discrimination.: Effects of brainstem infarcts and multiple sclerosis lesions. Hearing Research. 2000;143(1–2):29-42.
126. Levine RA, Gardner JC, Stufflebeam SM, Fullerton BC, Carlisle EW, Furst M, et al. Binaural auditory processing in multiple sclerosis subjects. Hearing Research. 1993;68(1):59-72.
127. Levine RA, Gardner JC, Fullerton BC, Stufflebeam SM, Furst M, Rosen BR. Multiple sclerosis lesions of the auditory pons are not silent. Brain. 1994;117(5):1127-41.
128. Hausler R, Levine RA. Brain stem auditory evoked potentials are related to interaural time discrimination in patients with multiple sclerosis. Brain Research. 1980;191(2):589-94.
129. Van Der Poel JC, Jones SJ, Miller DH. Sound lateralization, brainstem auditory evoked potentials and magnetic resonance imaging in multiple sclerosis. Brain. 1988;111(6):1453-74.
130. Lewis MS. Some effects of multiple sclerosis on speech perception in noise: preliminary findings. Journal of Rehabilitation Research & Development. 2006;43:xiii.
131. Goldstein DP, Stephens SD. Masking level difference: a measure of auditory processing capability. Audiology. 1975;14(4):354-67.
132. Hendler T, Squires NK, Emmerich DS. Psychophysical Measures of Central Auditory Dysfunction in Multiple Sclerosis: Neurophysiological and Neuroanatomical Correlates. Ear and Hearing. 1990;11(6):403-16.
133. Olsen WO, Noffsinger D, Carhart R. Masking level differences encountered in clinical populations. Audiology. 1976;15(4):287-301.
134. E. Musiek BF. The frequency pattern test: A guide. The Hearing Journal. 2002;55(6):58.
135. Mustek FE, Pinheiro ML. Frequency Patterns in Cochlear, Brainstem, and Cerebral Lesions: Reconnaissance mélodique dans les lésions cochléaires, bulbaires et corticales. Audiology. 1987;26(2):79-88.
136. Pantano P, Iannetti GD, Caramia F, Mainero C, Di Legge S, Bozzao L, et al. Cortical motor reorganization after a single clinical attack of multiple sclerosis. Brain. 2002;125(7):1607-15.
137. Prakash RS, Erickson KI, Snook EM, Colcombe SJ, Motl RW, Kramer AF. Cortical recruitment during selective attention in multiple sclerosis: An fMRI investigation of individual differences. Neuropsychologia. 2008;46(12):2888-95.
138. Musiek FE, Shinn JB, Jirsa R, Bamiou D-E, Baran JA, Zaida E. GIN (Gaps-In-Noise) Test Performance in Subjects with Confirmed Central Auditory Nervous System Involvement. Ear and Hearing. 2005;26(6):608-18.
139. Wong PCM, Uppunda AK, Parrish TB, Dhar S. Cortical mechanisms of speech perception in noise. Journal of Speech, Language, and Hearing Research. 2008;51:1026+.
140. Peters RW, Moore BCJ, Baer T. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. The Journal of the Acoustical Society of America. 1998;103(1):577-87.
141. Dunlop WA, Enticott PG, Rajan R. Speech discrimination difficulties in High-Functioning Autism Spectrum Disorder are likely independent of auditory hypersensitivity. Frontiers in Human Neuroscience. 2016;10.
142. Olsen WO, Noffsinger D, Kurdziel S. Speech Discrimination In Quiet And In White Noise By Patients With Peripheral And Central Lesions. Acta Oto-Laryngologica. 1975;80(1-6):375-82.

143. Dayal VS, Tarantino L, Swisher LP. Neuro-otologic studies in multiple sclerosis. The Laryngoscope. 1966;76(11):1798-809.
144. LeZak RJ, Selhub S. On hearing in multiple sclerosis. The Annals of otology, rhinology, and laryngology. 1966;75(4):1102-10.
145. Jacobson JT, Deppe U, Murray TJ. Dichotic Paradigms in Multiple Sclerosis. Ear & Hearing November/December. 1983;4(6):311-7.
146. Grenman R, Lang H, Panelius M, Salmivalli A, Laine H, Rintamaki J. Stapedius reflex and brainstem auditory evoked responses in multiple sclerosis patients. Scand Audiol. 1984;13(2):109-13.
147. Quine DB, Regan D, Murray TJ. Degraded Discrimination Between Speech-Like Sounds by Patients with Multiple Sclerosis and Friedreich's Ataxia. Brain. 1984;107(4):1113-22.
148. Quine DB, Regan D, Beverley KI, Murray TJ. PAtients with multiple sclerosis experience hearing loss specifically for shifts of tone frequency. Archives of Neurology. 1984;41(5):506-8.
149. Kimura D. Functional Asymmetry of the Brain in Dichotic Listening. Cortex. 1967;3(2):163-78.
150. Simon JH, Holtas SL, Schiffer RB, Rudick RA, Herndon RM, Kido DK, et al. Corpus callosum and subcallosal-periventricular lesions in multiple sclerosis: detection with MR. Radiology. 1986;160(2):363-7.
151. Westerhausen R, Hugdahl K. The corpus callosum in dichotic listening studies of hemispheric asymmetry: A review of clinical and experimental evidence. Neuroscience & Biobehavioral Reviews. 2008;32(5):1044-54.
152. Westerhausen R, Woerner W, Kreuder F, Schweiger E, Hugdahl K, Wittling W. The role of the corpus callosum in dichotic listening: a combined morphological and diffusion tensor imaging study. Neuropsychology. 2006;20(3):272-9.
153. Gadea M, Marti-Bonmati L, Arana E, Espert R, Casanova V, Pascual A. Dichotic Listening and Corpus Callosum Magnetic Resonance Imaging in Relapsing-Remitting Multiple Sclerosis With Emphasis on Sex Differences. Neuropsychology. 2002;16(2):275-81.
154. Barkhof F, Tas MW, Valk J, Elton M, Lindeboom J, Schmidt WF, et al. Functional correlates of callosal atrophy in relapsing-remitting multiple sclerosis patients. A preliminary MRI study. Journal of Neurology. 1998;245(3):153-8.
155. Kimura D. Cerebral cominance and the perception of verbal stimuli. Canadian Journal of Psychology. 1967;15:166-71.
156. Rubens AB, Froehling B, Slater G, Anderson D. Left ear suppression on verbal dichotic tests in patients with multiple sclerosis. Annals of Neurology. 1985;18(4):459-63.
157. Lindeboom J, ter Horst R. Interhemispheric disconnection effects in multiple sclerosis. Journal of Neurology, Neurosurgery & Psychiatry. 1988;51(11):1445.
158. Wishart HA, Strauss E, Hunter M, Moll A. Interhemispheric transfer in multiple sclerosis. Journal of Clinical and Experimental Neuropsychology. 1995;17(6):937-40.
159. Berlow YA, Pollaro J, Krisky C, Grinstead J, Snodgrass M, Kohout K, et al. Magnetic Resonance Imaging Correlates of Dichotic Listening Performance in Multiple Sclerosis. Semin Hear. 2012;33(03):283-94.

160. Valadbeigi A, Weisi F, Rohbakhsh N, Rezaei M, Heidari A, Rasa AR. Central auditory processing and word discrimination in patients with multiple sclerosis. European Archives of Oto-Rhino-Laryngology. 2014;271(11):2891-6.
161. Sperry JL, Wiley TL, Chial MR. Word recognition performance in various background competitors. Journal of the American Academy of Audiology. 1997;8(2):71-80.
162. Samuel AG. Speech Perception. Annual Review of Psychology. 2011;62(1):49-72.
163. Anderson S, Kraus N. Objective neural indices of speech-in-noise perception. Trends Amplif. 2010;14(2):73-83.
164. Anderson S, Skoe E, Chandrasekaran B, Kraus N. Neural timing is linked to speech perception in noise. J Neurosci. 2010;30(14):4922-6.
165. Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. Ear and hearing. 2010;31(3):302-24.
166. Musiek EF. Neuroanatomy, Neurophysiology, and Central Auditory Assessment. Part III: Corpus Callosum and Efferent Pathways. Ear and Hearing. 1986;7(6):349-58.
167. Amaral A, Silva C, Alves A, Prata C, Fernandes C, Simões P, et al. Associations between quality of life and central auditory processing in seniors: preliminary results. European Journal of Public Health. 2019;29(Supplement_1).
168. Folmer R, Hutter M, Lilly D, Shannon J, Casiana L, Wilmington D, et al. Electrophysiological Measures of Auditory Processing in Patients with Multiple Sclerosis. Seminars in Hearing. 2012;33(3):274-82.
169. Kohlrausch A, Braasch J, Kolossa D, Blauert J. An Introduction to Binaural Processing. In: Blauert J, editor. The Technology of Binaural Listening. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 1-32.
170. Hartelius L, Runmarker B, Andersen O, Nord L. Temporal Speech Characteristics of Individuals with Multiple Sclerosis and Ataxic Dysarthria: 'Scanning Speech' Revisited. Folia Phoniatrica et Logopaedica. 2000;52(5):228-38.
171. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-52.
172. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001;50(1):121-7.
173. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. International journal of surgery (London, England). 2014;12(12):1495-9.
174. Jerger JF, Carhart R, Tillman TW, Peterson JL. Some Relations Between Normal Hearing for Pure Tones and for Speech. Journal of Speech, Language, and Hearing Research. 1959;2(2):126-40.
175. Mann C, Canny B, Reser D, Rajan R. Poorer verbal working memory for a second language selectively impacts academic achievement in university medical students. PeerJ. 2013;1(1).

176. Bench J, Kowal A, Bamford J. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. British journal of audiology. 1979;13(3):108-12.
177. Sussman ES, Horváth J, Winkler I, Orr M. The role of attention in the formation of auditory streams. Perception & Psychophysics. 2007;69(1):136-52.
178. Nair PG, Basheer BM. Influence of temporal resolution skills in speech discrimination abilities of older subjects. Acta Otorhinolaryngol Ital. 2017;37(1):58-62.
179. Li N, Loizou PC. Factors influencing glimpsing of speech in noise. The Journal of the Acoustical Society of America. 2007;122(2):1165-72.
180. Bronkhorst A. The Cocktail Party Phenomenon: A Review of Research on Speech Intelligibility in Multiple-Talker Conditions. Acta Acustica united with Acustica. 2000;86:117-28.
181. Sur S, Sinha VK. Event-related potential: An overview. Ind Psychiatry J. 2009;18(1):70-3.
182. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. The Lancet Neurology. 2008;7(12):1139-51.
183. Wu C, Stefanescu RA, Martel DT, Shore SE. Listening to another sense: somatosensory integration in the auditory system. Cell and tissue research. 2015;361(1):233-50.
184. Correale J, Gaitán MI, Ysrraelit MC, Fiol MP. Progressive multiple sclerosis: from pathogenic mechanisms to treatment. Brain. 2016;140(3):527-46.
185. Wallin MT, Culpepper WJ, Nichols E, Bhutta ZA, Gebrehiwot TT, Hay SI, et al. Global, regional, and national burden of multiple sclerosis 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Neurology. 2019;18(3):269-85.
186. Cinar BP, Yorgun YG. What We Learned from The History of Multiple Sclerosis Measurement: Expanded Disability Status Scale. Noro Psikiyatr Ars. 2018;55(Suppl 1):S69-S75.
187. Amato MP, Ponziani G. Quantification of impairment in MS: discussion of the scales in use. Mult Scler. 1999;5(4):216-9.
188. Chiaravalloti ND, DeLuca J. Cognitive impairment in multiple sclerosis. Lancet Neurol. 2008;7(12):1139-51.
189. McNicholas N, O'Connell K, Yap SM, Killeen RP, Hutchinson M, McGuigan C. Cognitive dysfunction in early multiple sclerosis: a review. QJM: An International Journal of Medicine. 2017;111(6):359-64.
190. Lublin FD, Baier M, Cutter G. Effect of relapses on development of residual deficit in multiple sclerosis. Neurology. 2003;61(11):1528-32.
191. LoPresti P. Silent Free Fall at Disease Onset: A Perspective on Therapeutics for Progressive Multiple Sclerosis. Front Neurol. 2018;9(973):973.
192. Iva P, Fielding, J., Clough, M., White, O., Noffs, G., Godic, B., Martin, R., van der Walt, A., Rajan, R., . Speech Discrimination in Noise in Multiple Sclerosis Dataset. Data in Brief. 2020;(In Press).
193. Habek M. Evaluation of brainstem involvement in multiple sclerosis. Expert Rev Neurother. 2013;13(3):299-311.

194. Miller D, Barkhof F, Montalban X, Thompson A, Filippi M. Clinically isolated syndromes suggestive of multiple sclerosis, part I: natural history, pathogenesis, diagnosis, and prognosis. Lancet Neurol. 2005;4(5):281-8.
195. White-Schwoch T, Davies EC, Thompson EC, Woodruff Carr K, Nicol T, Bradlow AR, et al. Auditory-neurophysiological responses to speech during early childhood: Effects of background noise. Hear Res. 2015;328:34-47.
196. Dryden A, Allen HA, Henshaw H, Heinrich A. The Association Between Cognitive Performance and Speech-in-Noise Perception for Adult Listeners: A Systematic Literature Review and Meta-Analysis. Trends Hear. 2017;21:2331216517744675.
197. Kathleen Pichora-Fuller M. Use of supportive context by younger and older adult listeners: balancing bottom-up and top-down information processing. International journal of audiology. 2008;47 Suppl 2(sup2):S72-82.
198. Ronnberg J, Lunner T, Zekveld A, Sorqvist P, Danielsson H, Lyxell B, et al. The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. Front Syst Neurosci. 2013;7(31):31.
199. Stenfelt S, Ronnberg J. The signal-cognition interface: interactions between degraded auditory signals and cognitive processes. Scand J Psychol. 2009;50(5):385-93.
200. Zekveld AA, Rudner M, Johnsrude IS, Ronnberg J. The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. J Acoust Soc Am. 2013;134(3):2225-34.
201. Mattys SL, Davis MH, Bradlow AR, Scott SK. Speech recognition in adverse conditions: A review. Language and Cognitive Processes: Speech Recognition in Adverse Conditions. 2012;27(7-8):953-78.
202. Gronwall D.M. A. SHD. The psychological effects of concussion: Auckland : Auckland University Press ; Wellington : Oxford University Press 1974.
203. Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain. 1999;122 (Pt 5)(5):871-82.
204. Strober L, Deluca J, Benedict RH, Jacobs A, Cohen JA, Chiaravalloti N, et al. Symbol Digit Modalities Test: A valid clinical trial endpoint for measuring cognition in multiple sclerosis. Multiple sclerosis. 2019;25(13):1781-90.
205. Coo H, Hopman WM, Edgar CM, McBride EV, Brunet DG. The Paced Auditory Serial Addition Test: to what extent is it performed as instructed, and is it associated with disease course? Mult Scler. 2005;11(1):85-9.
206. Tombaugh TN. A comprehensive review of the Paced Auditory Serial Addition Test (PASAT). Arch Clin Neuropsychol. 2006;21(1):53-76.
207. Fisk JD, & Archibald, C. J. . Limitations of the paced auditory serial addition test as a measure of working memory in patients with multiple sclerosis. Journal of the International Neuropsychological Society. 2001;7(3): 363-72.
208. Mathias CW, Stanford MS, Houston RJ. The physiological experience of the Paced Auditory Serial Addition Task (PASAT): does the PASAT induce autonomic arousal? Arch Clin Neuropsychol. 2004;19(4):543-54.
209. Landrø NI, Sletvold H, Celius EG. Memory Functioning and Emotional Changes in Early Phase Multiple Sclerosis. Archives of Clinical Neuropsychology. 2000;15(1):37-46.
210. Mattys SL, Brooks J, Cooke M. Recognizing speech under a processing load: dissociating energetic from informational factors. Cogn Psychol. 2009;59(3):203-43.

211. Akeroyd MA. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. International journal of audiology. 2008;47 Suppl 2:S53-71.
212. Cherry EC. Some Experiments on the Recognition of Speech, with One and with Two Ears. The Journal of the Acoustical Society of America. 1953;25(5):975-9.
213. Kidd G, Colburn HS. Informational Masking in Speech Recognition. In: Middlebrooks JC, Simon JZ, Popper AN, Fay RR, editors. The Auditory System at the Cocktail Party. Springer Handbook of Auditory Research. Cham: Springer International Publishing; 2017. p. 75-109.
214. Rahn K, Slusher B, Kaplin A. Cognitive impairment in multiple sclerosis: a forgotten disability remembered. Cerebrum. 2012;2012:14.
215. Migliore S, Ghazaryan A, Simonelli I, Pasqualetti P, Squitieri F, Curcio G, et al. Cognitive Impairment in Relapsing-Remitting Multiple Sclerosis Patients with Very Mild Clinical Disability. Behav Neurol. 2017;2017:7404289.
216. Beck AS, RA. Manual for the revised Beck Depression Inventory. San Antonio, TX: The Psychological Corporation. 1987.
217. Nunnari D, De Cola MC, D'Aleo G, Rifici C, Russo M, Sessa E, et al. Impact of Depression, Fatigue, and Global Measure of Cortical Volume on Cognitive Impairment in Multiple Sclerosis. BioMed Research International. 2015;2015:519785.
218. Nelson H, Willison J. The revised national adult reading test–test manual. Windsor, UK: NFER-Nelson. 1991;991:1-6.
219. Fisk JD, Ritvo PG, Ross L, Haase DA, Marrie TJ, Schlech WF. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 1994;18 Suppl 1:S79-83.
220. Gronwall DM. Paced auditory serial-addition task: a measure of recovery from concussion. Perceptual and motor skills. 1977;44(2):367-73.
221. Smith A. Symbol Digit Modalities Test (Revised). Los Angeles: Western Psychological Services. 1982.
222. Delis DC, Kramer, J. H., Kaplan, E., & Ober, B. A. The California verbal learning test: Research edition, adult version. San Antonio, TX: The Psychological Corporation. 1987.
223. Weber H, Pfadenhauer K, Stöhr M, Rösler A. Central hyperacusis with phonophobia in multiple sclerosis. Multiple Sclerosis Journal. 2002;8(6):505-9.
224. Heinrich A, Knight S, editors. The Contribution of Auditory and Cognitive Factors to Intelligibility of Words and Sentences in Noise. Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing; 2016 2016//; Cham: Springer International Publishing.
225. Hosmer DW. Applied logistic regression. 2nd ed. ed. Lemeshow S, John W, Sons, Wiley I, editors. New York: New York : Wiley; 2000.
226. Cree BAC, Hollenbach JA, Bove R, Kirkish G, Sacco S, Caverzasi E, et al. Silent progression in disease activity-free relapsing multiple sclerosis. Ann Neurol. 2019;85(5):653-66.

227. Brochet B, Ruet A. Cognitive Impairment in Multiple Sclerosis With Regards to Disease Duration and Clinical Phenotypes. Frontiers in Neurology. 2019;10.
228. Lynch SG, Parmenter BA, Denney DR. The association between cognitive impairment and physical disability in multiple sclerosis. Mult Scler. 2005;11(4):469-76.
229. Sheldrake J, Diehl PU, Schaette R. Audiometric characteristics of hyperacusis patients. Frontiers in neurology. 2015;6:105-.
230. Sussman ES. Auditory Scene Analysis: An Attention Perspective. Journal of speech, language, and hearing research : JSLHR. 2017;60(10):2989-3000.
231. Janse E. A non-auditory measure of interference predicts distraction by competing speech in older adults. Aging, Neuropsychology, and Cognition. 2012;19(6):741-58.
232. Simpson SA, Cooke M. Consonant identification in N-talker babble is a nonmonotonic function of N. The Journal of the Acoustical Society of America. 2005;118(5):2775-8.
233. Pressnitzer D, Suied C, Shamma S. Auditory scene analysis: The sweet music of ambiguity. Frontiers in Human Neuroscience. 2011;5(158).
234. Füllgrabe C, Rosen S. Investigating the Role of Working Memory in Speech-in-noise Identification for Listeners with Normal Hearing. Advances in experimental medicine and biology. 2016;894:29-36.
235. Sreenivasan KK, Sambhara D, Jha AP. Working memory templates are maintained as feature-specific perceptual codes. Journal of neurophysiology. 2011;106(1):115-21.
236. Costa SL, Genova HM, Deluca J, Chiaravalloti ND. Information processing speed in multiple sclerosis: Past, present, and future. Multiple Sclerosis Journal. 2017;23(6):772-89.
237. Drew MA, Starkey NJ, Isler RB. Examining the Link between Information Processing Speed and Executive Functioning in Multiple Sclerosis. Archives of Clinical Neuropsychology. 2009;24(1):47-58.
238. Forn C, Belenguer A, Parcet-Ibars MA, Ávila C. Information-processing speed is the primary deficit underlying the poor performance of multiple sclerosis patients in the Paced Auditory Serial Addition Test (PASAT). Journal of Clinical and Experimental Neuropsychology. 2008;30(7):789-96.
239. Benedict RH, DeLuca J, Phillips G, LaRocca N, Hudson LD, Rudick R, et al. Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. Multiple sclerosis (Houndmills, Basingstoke, England). 2017;23(5):721-33.
240. Amato MP, Ponziani G, Pracucci G, Bracco L, Siracusa G, Amaducci L. Cognitive Impairment in Early-Onset Multiple Sclerosis: Pattern, Predictors, and Impact on Everyday Life in a 4-Year Follow-up. Archives of Neurology. 1995;52(2):168-72.
241. Cullum CM. 4.11 - Neuropsychological Assessment of Adults. In: Bellack AS, Hersen M, editors. Comprehensive Clinical Psychology. Oxford: Pergamon; 1998. p. 303-47.
242. St Clair-Thompson HL, Allen RJ. Are forward and backward recall the same? A dual-task study of digit recall. Memory & Cognition. 2013;41(4):519-32.

243. Avivi-Reich M, Daneman M, Schneider B. How age and linguistic competence alter the interplay of perceptual and cognitive factors when listening to conversations in a noisy environment. Frontiers in Systems Neuroscience. 2014;8(21).
244. Pisoni DB. Word Identification in Noise. Lang Cogn Process. 1996;11(6):681-7.
245. Musiek FE, Chermak GD. Chapter 18 - Psychophysical and behavioral peripheral and central auditory tests. In: Michael J. Aminoff FB, Dick FS, editors. Handbook of Clinical Neurology. Volume 129: Elsevier; 2015. p. 313-32.
246. Musiek FE, Baran JA, Shinn JB. Disorders of the Auditory System. San Diego, UNITED STATES: Plural Publishing, Inc.; 2011.
247. Yin TCT. Neural Mechanisms of Encoding Binaural Localization Cues in the Auditory Brainstem. In: Oertel D, Fay RR, Popper AN, editors. Integrative Functions in the Mammalian Auditory Pathway. New York, NY: Springer New York; 2002. p. 99-159.
248. Bacon JH, Kister I, Bacon TE, Pasternak E, Strauchler Y, Herbert J. Sound Lateralization Test Distinguishes Unimpaired MS Patients from Healthy Controls. Multiple Sclerosis International. 2014;2014:462043.
249. Freyman RL, Helfer KS, McCall DD, Clifton RK. The role of perceived spatial separation in the unmasking of speech. The Journal of the Acoustical Society of America. 1999;106(6):3578-88.
250. Bronkhorst AW. The cocktail-party problem revisited: early processing and selection of multi-talker speech. Attention, perception & psychophysics. 2015;77(5):1465.
251. Litovsky RY. Spatial Release From Masking. Acoustics Today. 2012;8(2):7.
252. Plomp R, Mimpen AM. Effect of the Orientation of the Speaker's Head and the Azimuth of a Noise Source on the Speech-Reception Threshold for Sentences. Acta Acustica united with Acustica. 1981;48(5):325-8.
253. Yost WA. Spatial release from masking based on binaural processing for up to six maskers. The Journal of the Acoustical Society of America. 2017;141(3):2093-.
254. Freyman RL, Balakrishnan U, Helfer KS. Spatial release from informational masking in speech recognition. J Acoust Soc Am. 2001;109(5 Pt 1):2112-22.
255. Culling JF, Hawley ML, Litovsky RY. The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources. The Journal of the Acoustical Society of America. 2004;116(2):1057-65.
256. Glyde H, Hickson L, Cameron S, Dillon H. Problems hearing in noise in older adults: a review of spatial processing disorder. Trends Amplif. 2011;15(3):116-26.
257. Bronkhorst AW, Plomp R. The effect of head-induced interaural time and level differences on speech intelligibility in noise. The Journal of the Acoustical Society of America. 1988;83(4):1508.
258. Hawley ML, Litovsky RY, Culling JF. The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. The Journal of the Acoustical Society of America. 2004;115(2):833.
259. Kolarik AJ, Moore BCJ, Zahorik P, Cirstea S, Pardhan S. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. Attention, perception & psychophysics. 2016;78(2):373-95.

260. Popescu BFG, Pirko I, Lucchinetti CF. Pathology of multiple sclerosis: where do we stand? Continuum (Minneap Minn). 2013;19(4 Multiple Sclerosis):901-21.
261. Moore DR. Anatomy and physiology of binaural hearing. Audiology. 1991;30(3):125-34.
262. Strange W, Workshop on Cross-Language Speech P, editors. Speech perception and linguistic experience : issues in cross-language research1995; Baltimore: York Press.
263. van Wijngaarden SJ, Steeneken HJM, Houtgast T. Quantifying the intelligibility of speech in noise for non-native listeners. The Journal of the Acoustical Society of America. 2002;111(4):1906-16.
264. Bradlow AR, Bent T. The clear speech effect for non-native listeners. The Journal of the Acoustical Society of America. 2002;112(1):272-84.
265. Cainer KE, James C, Rajan R. Learning speech-in-noise discrimination in adult humans. Hearing Research. 2008;238(1):155-64.
266. Martin RLM, Ken I.; Senova, Melis A. Free-Field Equivalent Localization of Virtual Audio. JAES. 2001;49(1/2):14-22.
267. Arbogast TL, Mason CR, Kidd G. The effect of spatial separation on informational and energetic masking of speech. The Journal of the Acoustical Society of America. 2002;112(5 Pt 1):2086-98.
268. Srinivasan NK, Jakien KM, Gallun FJ. Release from masking for small spatial separations: Effects of age and hearing loss. The Journal of the Acoustical Society of America. 2016;140(1):EL73-EL.
269. Samson F, Johnsrude I. Effects of a consistent target or masker voice on target speech intelligibility in two- and three-talker mixtures. The Journal of the Acoustical Society of America. 2016;139:1037-46.
270. Brungart D, Chang P, Simpson B, Wang D. Isolating the Energetic Component of Speech-On-Speech Masking with Ideal Time-Frequency Segregation. The Journal of the Acoustical Society of America. 2006;120(6):4007-18.
271. Domingo Y, Holmes E, Macpherson E, Johnsrude IS. Using spatial release from masking to estimate the magnitude of the familiar-voice intelligibility benefit. The Journal of the Acoustical Society of America. 2019;146(5):3487-94.
272. Brungart DS, Simpson BD, Ericson MA, Scott KR. Informational and energetic masking effects in the perception of multiple simultaneous talkers. The Journal of the Acoustical Society of America. 2001;110(5):2527-38.
273. Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. The Journal of the Acoustical Society of America. 2001;109(3):1101-9.
274. Shinn-Cunningham B. Influences of spatial cues on grouping and understanding sound. Proc Forum Acusticum. 2005.
275. Schwartz A, McDermott JH, Shinn-Cunningham B. Spatial cues alone produce inaccurate sound segregation: the effect of interaural time differences. The Journal of the Acoustical Society of America. 2012;132(1):357-68.
276. Zekveld A, Rudner M, Kramer S, Lyzenga J, Rönnberg J. Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. Frontiers in Neuroscience. 2014;8(88).

277. Xia J, Nooraei N, Kalluri S, Edwards B. Spatial release of cognitive load measured in a dual-task paradigm in normal-hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America. 2015;137(4):1888.
278. Andéol G, Suied C, Scannella S, Dehais F. The Spatial Release of Cognitive Load in Cocktail Party Is Determined by the Relative Levels of the Talkers. J Assoc Res Otolaryngol. 2017;18(3):457-64.
279. Amato MP, Ponziani G, Siracusa G, Sorbi S. Cognitive Dysfunction in Early-Onset Multiple Sclerosis: A Reappraisal After 10 Years. Archives of Neurology. 2001;58(10):1602-6.
280. Vanthornhout J, Decruy L, Francart T. Effect of Task and Attention on Neural Tracking of Speech. Frontiers in Neuroscience. 2019;13(977).
281. Posner MI, Petersen SE. The Attention System of the Human Brain. Annual Review of Neuroscience. 1990;13(1):25-42.
282. Petersen SE, Posner MI. The attention system of the human brain: 20 years after. Annual review of neuroscience. 2012;35:73-89.
283. Scott S, McGettigan C. The neural processing of masked speech. Hearing research. 2013;303.
284. Wild CJ, Yusuf A, Wilson DE, Peelle JE, Davis MH, Johnsrude IS. Effortful listening: the processing of degraded speech depends critically on attention. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2012;32(40):14010-21.
285. Romanski LM, Bates JF, Goldman-Rakic PS. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol. 1999;403(2):141-57.
286. Salmi J, Rinne T, Koistinen S, Salonen O, Alho K. Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. Brain Research. 2009;1286:155-64.
287. Scott SK, Rosen S, Wickham L, Wise RJ. A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. J Acoust Soc Am. 2004;115(2):813-21.
288. Paus T, Zatorre RJ, Hofle N, Caramanos Z, Gotman J, Petrides M, et al. Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. J Cogn Neurosci. 1997;9(3):392-408.
289. Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M. The Essential Role of Premotor Cortex in Speech Perception. Current Biology. 2007;17(19):1692-6.
290. Du Y, Buchsbaum BR, Grady CL, Alain C. Increased activity in frontal motor cortex compensates impaired speech perception in older adults. Nature Communications. 2016;7(1):12241.
291. Wong PCM, Jin JX, Gunasekera GM, Abel R, Lee ER, Dhar S. Aging and cortical mechanisms of speech perception in noise. Neuropsychologia. 2009;47(3):693-703.
292. Reuter-Lorenz P, Cappell K. Neurocognitive Aging and the Compensation Hypothesis. Current Directions in Psychological Science - CURR DIRECTIONS PSYCHOL SCI. 2008;17:177-82.
293. Horakova D, Kalincik T, Dusankova JB, Dolezal O. Clinical correlates of grey matter pathology in multiple sclerosis. BMC neurology. 2012;12:10-.

294. Eshaghi A, Marinescu RV, Young AL, Firth NC, Prados F, Jorge Cardoso M, et al. Progression of regional grey matter atrophy in multiple sclerosis. Brain. 2018;141(6):1665-77.
295. Chard DT, Miller DH. What lies beneath grey matter atrophy in multiple sclerosis? Brain. 2016;139(1):7-10.
296. Benedict RH, Weinstock-Guttman B, Fishman I, Sharma J, Tjoa CW, Bakshi R. Prediction of neuropsychological impairment in multiple sclerosis: comparison of conventional magnetic resonance imaging measures of atrophy and lesion burden. Arch Neurol. 2004;61(2):226-30.
297. Houtchens MK, Benedict RH, Killiany R, Sharma J, Jaisani Z, Singh B, et al. Thalamic atrophy and cognition in multiple sclerosis. Neurology. 2007;69(12):1213-23.
298. Enzinger C, Fazekas F. Measuring Gray Matter and White Matter Damage in MS: Why This is Not Enough. Frontiers in neurology. 2015;6:56-.
299. Metere R, Kober T, Möller HE, Schäfer A. Simultaneous Quantitative MRI Mapping of T1, T2* and Magnetic Susceptibility with Multi-Echo MP2RAGE. PLoS One. 2017;12(1):e0169265.
300. Rojas JI, Patrucco L, Miguez J, Cristiano E. Brain atrophy in multiple sclerosis: therapeutic, cognitive and clinical impact. Arquivos de Neuro-Psiquiatria. 2016;74:235-43.
301. Hickok G. The functional neuroanatomy of language. Phys Life Rev. 2009;6(3):121-43.
302. Poeppel Ha. The Cortical Organisation of Speech Processing. Perspectives. 2007;8:393-402.
303. DeWitt I, Rauschecker JP. Phoneme and word recognition in the auditory ventral stream. Proceedings of the National Academy of Sciences. 2012;109(8):E505-E14.
304. Specht K. Neuronal basis of speech comprehension. Hearing Research. 2014;307:121-35.
305. de Heer WA, Huth AG, Griffiths TL, Gallant JL, Theunissen FE. The Hierarchical Cortical Organization of Human Speech Processing. The Journal of Neuroscience. 2017;37(27):6539.
306. Wong PCM, Ettlinger M, Sheppard JP, Gunasekera GM, Dhar S. Neuroanatomical characteristics and speech perception in noise in older adults. Ear and hearing. 2010;31(4):471-9.
307. Štecková T, Hluštík P, Sládková V, Odstrčil F, Mareš J, Kaňovský P. Thalamic atrophy and cognitive impairment in clinically isolated syndrome and multiple sclerosis. Journal of the neurological sciences. 2014;342(1-2):62.
308. Morgen K, Sammer G, Courtney SM, Wolters T, Melchior H, Blecker CR, et al. Evidence for a direct association between cortical atrophy and cognitive impairment in relapsing–remitting MS. NeuroImage. 2006;30(3):891-8.
309. Matias-Guiu JA, Cortés-Martínez A, Montero P, Pytel V, Moreno-Ramos T, Jorquera M, et al. Structural MRI correlates of PASAT performance in multiple sclerosis. BMC Neurology. 2018;18(1):214.
310. Nocentini U, Bozzali M, Spanò B, Cercignani M, Serra L, Basile B, et al. Exploration of the relationships between regional grey matter atrophy and cognition in multiple sclerosis. Brain Imaging and Behavior. 2014;8(3):378-86.
311. Snyder PJ, CappelJeri JC. Information processing speed deficits may be better correlated with the extent of white matter sclerotic lesions in multiple sclerosis than previously suspected. Brain and Cognition. 2001;46(1):279-84.

312. Manca R, Sharrack B, Paling D, Wilkinson ID, Venneri A. Brain connectivity and cognitive processing speed in multiple sclerosis: A systematic review. Journal of the Neurological Sciences. 2018;388:115-27.
313. Baltruschat SA, Ventura-Campos N, Cruz-Gómez ÁJ, Belenguer A, Forn C. Gray matter atrophy is associated with functional connectivity reorganization during the Paced Auditory Serial Addition Test (PASAT) execution in Multiple Sclerosis (MS). Journal of Neuroradiology. 2015;42(3):141-9.
314. Mahad DH, Trapp BD, Lassmann H. Pathological mechanisms in progressive multiple sclerosis. The Lancet Neurology. 2015;14(2):183-93.
315. Koriem KMM. Multiple sclerosis: New insights and trends. Asian Pacific Journal of Tropical Biomedicine. 2016;6(5):429-40.
316. Nachtegaal J, Festen JM, Kramer SE. Hearing Ability and its Relationship with Psychosocial Health, Work-Related Variables, and Health Care Use: The National Longitudinal Study on Hearing. Audiol Res. 2011;1(1):e9-e.
317. Bregman AS. Auditory scene analysis: The perceptual organization of sound. Auditory scene analysis: The perceptual organization of sound. Cambridge, MA, US: The MIT Press; 1990. p. xiii, 773-xiii, .
318. Zivadinov R, Sepcic J, Nasuelli D, De Masi R, Bragadin LM, Tommasi MA, et al. A longitudinal study of brain atrophy and cognitive disturbances in the early phase of relapsing-remitting multiple sclerosis. Journal of Neurology, Neurosurgery & Psychiatry2001.
319. Ferreira MLB. Cognitive deficits in multiple sclerosis: a systematic review. Arquivos de Neuro-Psiquiatria. 2010;68:632-41.
320. Nogueira LAC, Teixeira L, Sabino P, Filho HA, Alvarenga RMP, Thuler LC. Gait characteristics of multiple sclerosis patients in the absence of clinical disability. Disability and rehabilitation. 2013;35(17):1472-8.
321. Novotna K, Sobisek L, Horakova D, Havrdova E, Lizrova Preiningerova J. Quantification of Gait Abnormalities in Healthy-Looking Multiple Sclerosis Patients (with Expanded Disability Status Scale 0-1.5). European neurology. 2016;76(3-4):99-104.
322. Migliore S, Ghazaryan A, Simonelli I, Pasqualetti P, Squitieri F, Curcio G, et al. Cognitive Impairment in Relapsing-Remitting Multiple Sclerosis Patients with Very Mild Clinical Disability.(Research Article). Behavioural Neurology. 2017;2017.
323. Guzel I, Mungan S, Oztekin ZN, Ak F. Is there an association between the Expanded Disability Status Scale and inflammatory markers in multiple sclerosis? Journal of the Chinese Medical Association. 2016;79(2):54-7.
324. Expanded Disability Status Scale: Multiple Sclerosis Trust; 2020 [Available from: https://mstrust.org.uk/a-z/expanded-disability-status-scale-edss#:~:text=The%20EDSS%20scale%20ranges%20from,eight%20functional%20systems%20(FS)%3A.
325. Functional Systems Scores (FSS) and Expanded Disability Status Scale (EDSS): National Multiple Sclerosis Society; 2020 [Available from: https://www.nationalmssociety.org/NationalMSSociety/media/MSNationalFiles/Brochures/10-2-3-29-EDSS_Form.pdf.
326. Şen S. Neurostatus and EDSS Calculation with Cases. Noro psikiyatri arsivi. 2018;55(Suppl 1):S80-S3.

327. Coles AJ, Cox A, Le Page E, Jones J, Trip SA, Deans J, et al. The window of therapeutic opportunity in multiple sclerosis. Journal of Neurology. 2006;253(1):98-108.
328. Mathes T, Pieper D. Clarifying the distinction between case series and cohort studies in systematic reviews of comparative studies: potential impact on body of evidence and workload. BMC Med Res Methodol. 2017;17(1):107-.
329. Jongen PJ. Observational designs in clinical multiple sclerosis research: Particulars, practices and potentialities. Multiple Sclerosis and Related Disorders. 2019;35:142-9.
330. Fulton SE, Lister JJ, Bush ALH, Edwards JD, Andel R. Mechanisms of the Hearing-Cognition Relationship. Seminars in hearing. 2015;36(3):140-9.
331. Lee K-Y. Pathophysiology of age-related hearing loss (peripheral and central). Korean J Audiol. 2013;17(2):45-9.
332. Peyvandi AN, Bijan; Roozbahany, Navid Ahmady. Neuro-otologic Manifestations of Multiple Sclerosis. Archives of Iranian Medicine. 2010;188-92.
333. Nixon GK, Sarant JZ, Tomlin D. Peripheral and central hearing impairment and their relationship with cognition: a review. International journal of audiology. 2019;58(9):541-52.
334. Fortunato S, Forli F, Guglielmi V, De Corso E, Paludetti G, Berrettini S, et al. A review of new insights on the association between hearing loss and cognitive decline in ageing. Acta Otorhinolaryngol Ital. 2016;36(3):155-66.
335. Mosnier I, Vanier A, Bonnard D, Lina-Granade G, Truy E, Bordure P, et al. Long-Term Cognitive Prognosis of Profoundly Deaf Older Adults After Hearing Rehabilitation Using Cochlear Implants. Journal of the American Geriatrics Society. 2018;66(8):1553.
336. Anderson S, Parbery-Clark A, Yi H-G, Kraus N. A neural basis of speech-in-noise perception in older adults. Ear and hearing. 2011;32(6):750-7.
337. Bajo VM, Nodal FR, Moore DR, King AJ. The descending corticocollicular pathway mediates learning-induced auditory plasticity. Nature neuroscience. 2010;13(2):253-60.
338. Bamiou DE, Musiek FE, Luxon LM. Aetiology and clinical presentations of auditory processing disorders--a review. Archives of disease in childhood. 2001;85(5):361-5.
339. Portnuff C, Bell B. Effective Use of Speech-in-Noise Testing in the Clinic. The Hearing Journal. 2019;72(5).
340. Humes LE, Dubno JR, Gordon-Salant S, Lister JJ, Cacace AT, Cruickshanks KJ, et al. Central presbycusis: a review and evaluation of the evidence. Journal of the American Academy of Audiology. 2012;23(8):635-66.
341. Tun PA, Williams VA, Small BJ, Hafter ER. The effects of aging on auditory processing and cognition. American journal of audiology. 2012;21(2):344-50.
342. Lee JS, Park WK, Kim WL-S, Kim WH. Effects of Noise Level and Cognitive Function on Speech Perception in Normal Elderly and Elderly with Amnestic Mild Cognitive Impairment. Cognitive And Behavioral Neurology. 2016;29(2):68-77.
343. Aimoni C, Prosser S, Ciorba A, Menozzi L, Soavi C, Zuliani G. Speech Audiometry Tests in Noise Are Impaired in Older Patients with Mild Cognitive Impairment: A Pilot Study: MJO. The Journal of International Advanced Otology. 2014;10(3):228-33.

344. Knopman DS, Petersen RC. Mild cognitive impairment and mild dementia: a clinical perspective. Mayo Clin Proc. 2014;89(10):1452-9.

345. Rajan R, Cainer KE. Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. Neuroscience. 2008;154(2):784-95.

346. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. The Lancet. 2018;392(10146):508-20.

347. Rosen S, Souza P, Ekelund C, Majeed AA. Listening to speech in a background of other talkers: effects of talker number and noise vocoding. The Journal of the Acoustical Society of America. 2013;133(4):2431-43.

348. Miller GA. The masking of speech. Psychological Bulletin. 1947;44(2):105-29.

349. Hervais-Adelman A, Pefkou M, Golestani N. Bilingual speech-in-noise: Neural bases of semantic context use in the native language. Brain and Language. 2014;132:1-6.

350. Brefczynski-Lewis JA, Lewis JW. Auditory object perception: A neurobiological model and prospective review. Neuropsychologia. 2017;105:223-42.

351. ten Oever S, Schroeder CE, Poeppel D, van Atteveldt N, Zion-Golumbic E. Rhythmicity and cross-modal temporal cues facilitate detection. Neuropsychologia. 2014;63:43-50.

352. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive psychology. 2000;41(1):49-100.

353. Diamond A. Executive functions. Annual review of psychology. 2013;64:135-68.

354. Peelle JE. Listening Effort: How the Cognitive Consequences of Acoustic Challenge Are Reflected in Brain and Behavior. Ear and hearing. 2018;39(2):204-14.

355. Giovannoni G, Butzkueven H, Dhib-Jalbut S, Hobart J, Kobelt G, Pepper G, et al. Brain health: time matters in multiple sclerosis. Mult Scler Relat Disord. 2016;9 Suppl 1:S5-s48.

356. Amato MP, Bartolozzi ML, Zipoli V, Portaccio E, Mortilla M, Guidi L, et al. Neocortical volume decrease in relapsing–remitting MS patients with mild cognitive impairment. Neurology. 2004;63(1):89-93.

357. Alain C, Du Y, Bernstein LJ, Barten T, Banai K. Listening under difficult conditions: An activation likelihood estimation meta-analysis. Human Brain Mapping. 2018;39(7):2695-709.

358. van Wijngaarden SJ, Steeneken HJ, Houtgast T. Quantifying the intelligibility of speech in noise for non-native talkers. J Acoust Soc Am. 2002;112(6):3004-13.

359. Feenaughty L, Tjaden K, Benedict RHB, Weinstock-Guttman B. Speech and pause characteristics in multiple sclerosis: a preliminary study of speakers with high and low neuropsychological test performance. Clin Linguist Phon. 2013;27(2):134-51.